Self-consistent Bulge/Disk/Halo Galaxy Dynamical Modeling Using Integral Field Kinematics

D. S. Taranu[1,2] , D. Obreschkow[1,2], J. J. Dubinski[3], L. M. R. Fogarty[2,4] , J. van de Sande[4] , B. Catinella[1] , L. Cortese[1] , A. Moffett[1], A. S. G. Robotham[1] , J. T. Allen[2,4] , J. Bland-Hawthorn[4] , J. J. Bryant[2,4,5], M. Colless[2,6] , S. M. Croom[2,4], F. D'Eugenio[6], R. L. Davies[7] , M. J. Drinkwater[2,8] , S. P. Driver[1] , M. Goodwin[5], I. S. Konstantopoulos[9], J. S. Lawrence[5], Á. R. López-Sánchez[5,10] , N. P. F. Lorente[5], A. M. Medling[6,11,14] , J. R. Mould[12] , M. S. Owers[5,10], C. Power[1,2] , S. N. Richards[2,4,5] , and C. Tonini[13]

[1] International Centre for Radio Astronomy Research, The University of Western Australia, 35 Stirling Highway, Crawley, WA, 6009, Australia
[2] ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Australia
[3] Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON, M5S 3H4, Canada
[4] Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW, 2006, Australia
[5] Australian Astronomical Observatory, PO Box 915, North Ryde, NSW, 1670, Australia
[6] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT, 2611, Australia
[7] Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK
[8] School of Mathematics and Physics, University of Queensland, QLD, 4072, Australia
[9] Envizi Suite 213, National Innovation Centre, Australian Technology Park, 4 Cornwallis Street, Eveleigh NSW 2015, Australia
[10] Department of Physics and Astronomy, Macquarie University, NSW, 2109, Australia
[11] Cahill Center for Astronomy and Astrophysics California Institute of Technology, MS 249-17, Pasadena, CA, 91125, USA
[12] Swinburne University, Hawthorn, VIC, 3122, Australia
[13] Melbourne University, School of Physics, Parkville, VIC, 3010, Australia

Abstract

We introduce a method for modeling disk galaxies designed to take full advantage of data from integral field spectroscopy (IFS). The method fits equilibrium models to simultaneously reproduce the surface brightness, rotation, and velocity dispersion profiles of a galaxy. The models are fully self-consistent 6D distribution functions for a galaxy with a Sérsic profile stellar bulge, exponential disk, and parametric dark-matter halo, generated by an updated version of GalactICS. By creating realistic flux-weighted maps of the kinematic moments (flux, mean velocity, and dispersion), we simultaneously fit photometric and spectroscopic data using both maximum-likelihood and Bayesian (MCMC) techniques. We apply the method to a GAMA spiral galaxy (G79635) with kinematics from the SAMI Galaxy Survey and deep g- and r-band photometry from the VST-KiDS survey, comparing parameter constraints with those from traditional 2D bulge–disk decomposition. Our method returns broadly consistent results for shared parameters while constraining the mass-to-light ratios of stellar components and reproducing the H I-inferred circular velocity well beyond the limits of the SAMI data. Although the method is tailored for fitting integral field kinematic data, it can use other dynamical constraints like central fiber dispersions and H I circular velocities, and is well-suited for modeling galaxies with a combination of deep imaging and H I and/or optical spectra (resolved or otherwise). Our implementation (MagRite) is computationally efficient and can generate well-resolved models and kinematic maps in under a minute on modern processors.

Key words: galaxies: fundamental parameters – galaxies: spiral – galaxies: structure – methods: data analysis

1. Introduction

The physical properties of nearby spiral galaxies are typically derived by fitting a number of distinct components to broadband images, either using azimuthally averaged 1D profiles or directly in 2D (e.g., Peng et al. 2002; Sánchez-Janssen et al. 2016; Johnston et al. 2017). For large surveys, common models are single Sérsic (1968) profile fits or two-component bulge–disk decompositions using an exponential disk and a de Vaucouleurs (1959) or Sérsic profile bulge (Simard et al. 2011)—appropriate parametrizations for galaxy disks and "classical" dispersion-supported bulges (Gadotti 2009), respectively. Additional features like bars, spiral arms, and dust are usually only modeled for well-resolved nearby galaxies.

Photometric bulge–disk decomposition has several major drawbacks. First, the best-fit 2D model may be impossible to reproduce with more realistic 3D density profiles or a 6D phase space distribution function (DF)—a serious concern, since most nearby galaxies are dynamically relaxed systems close to virial equilibrium. Two-dimensional models may be unable to produce a stable equilibrium system or require an unrealistic dark-matter halo density profile to reproduce the rotation curve. Therefore, it is desirable that fitting methods exclude parameter combinations that cannot create stable equilibrium models consistent with the galaxy's dynamics.

Bulge–disk decompositions can also produce ambiguous results. For fits with an exponential and a Sérsic profile, it is often assumed that the exponential component is a disk, whereas the Sérsic component is a bulge; however, the bulge can have a best-fit Sérsic index $n_s \approx 1$, leaving only the size and ellipticity to distinguish it from the disk. Furthermore, bulges are typically centrally concentrated and compact, and therefore difficult to resolve beyond $z > 0.05$ with seeing-limited imaging (Kelvin et al. 2014).

Kinematic data can break these degeneracies, even if spatially unresolved. A single-fiber central velocity dispersion can be used to infer the presence of a "classical," dispersion-
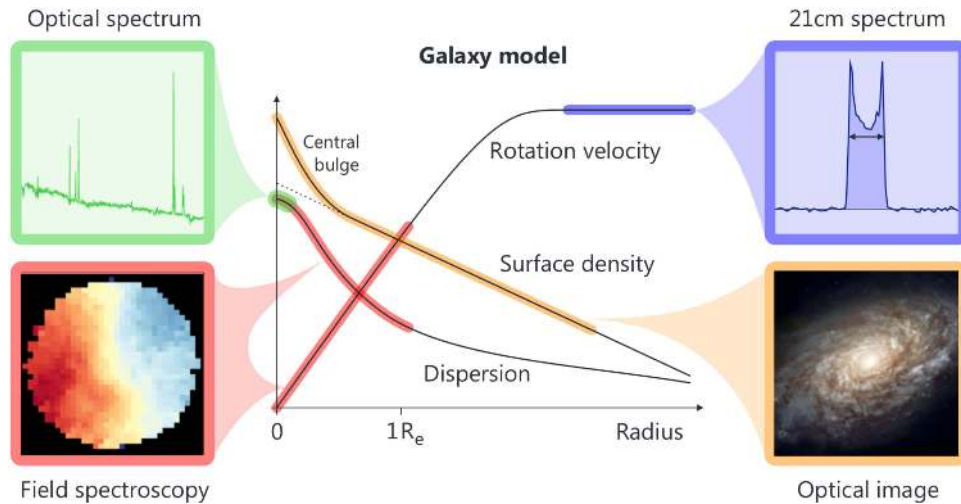
---

[14] Hubble Fellow.

Figure 1. Illustration of how various multiwavelength data can constrain disk galaxy dynamics over different regions of a typical massive spiral galaxy with an extended disk, concentrated bulge, and flat rotation curve.

supported bulge. Similarly, an unresolved H I 21 cm spectrum exhibiting a "double-horned" profile traces the orbital velocity of the neutral hydrogen gas, constraining the circular velocity at a large physical radius and (in principle) the combination of the rotation curve and H I surface density profile.

Integral field spectroscopy (IFS) permits the inference of spatially resolved rotation and dispersion profiles by taking multiple spectra across each galaxy. Single-target surveys that are already completed include SAURON (de Zeeuw et al. 2002), ATLAS3D (Cappellari et al. 2011), and CALIFA (Sánchez et al. 2012, 2016)—with nearly 1000 galaxies among them—whereas ongoing multiplexed surveys like SAMI (Croom et al. 2012; Bryant et al. 2015) and MaNGa (Bundy et al. 2015) have each observed ∼2000 galaxies and are expected to finish with 3600/∼10,000, respectively. These data sets have enormous potential to constrain fundamental galaxy properties, as illustrated in Figure 1, particularly for multi-component galaxies and when combined with multiwavelength data like deep imaging and H I spectra.

There are two major challenges in interpreting IFS kinematic maps. First, extracting information on galaxy kinematics requires careful modeling to account for observational and instrumental effects, particularly "beam smearing"—the tendency for a point-spread function (PSF) to blur ordered rotation across a galaxy, artificially increasing the velocity dispersion. Creating spectral datacubes by stacking dithered observations has an adverse impact on image resolution, particularly in the presence of differential atmospheric refraction (Law et al. 2015; Sharp et al. 2015), an issue that can and should be resolved by forward modeling rather than in the data reduction process. Second, IFS maps may not have sufficient spatial coverage or signal-to-noise ratio to reach the peak of a typical galactic rotation curve, whereas even unresolved 21 cm H I spectra can, since H I disks are typically more extended than stellar disks (e.g., Walter et al. 2008; Wang et al. 2016).

Our new modeling method is designed to resolve the issues outlined above. We create dynamical models from fully self-consistent phase space DFs, then generate synthetic observations of the kinematic moments to compare with observed data. Using kinematic moment maps allow for less ambiguous detections of dispersion-supported bulges. Synthetic observations reproduce biases from beam smearing through the PSF/

line-spread function (LSF) and pixel discretization, allowing for simultaneous fitting of independent data sets. Finally, since the models are based on theoretically motivated analytic density profiles, they predict reasonable extrapolations beyond the limits of the observed data—vital for estimating the angular momentum in extended disks (Romanowsky & Fall 2012; Obreschkow & Glazebrook 2014).

Existing galaxy dynamical modeling methods include Schwarzschild (1979) modeling (e.g., Cappellari et al. 2006), Jeans' modeling (as reviewed by Courteau et al. 2014), and made-to-measure (Syer & Tremaine 1996) and action-based modeling (e.g., Binney & McMillan 2011). However, most of these methods are not specifically designed to perform bulge–disk decomposition (but see Vasiliev & Athanassoula 2015) and many do not necessarily produce self-consistent DFs (see, e.g., Trick et al. 2016, who model the Milky Way's disk DF including a halo potential but no halo DF). Portail et al. (2017) fit Milky Way data using a near-equilibrium M2M model with a disk, halo, bulge, and bar, but at a significant computational cost of 190 CPU-hours for 25 iterations. Our method solves both problems, generating synthetic observations of near-equilibrium bulge/disk/halo models efficiently enough to fit data from large surveys like SAMI.

In Section 2, we describe the data sources for the sample galaxy used in this pilot study. In Section 3, we describe each step of the method in greater detail. In Section 4, we show more detailed results and comparisons to 2D bulge–disk decomposition, summarizing conclusions and outlining future directions in Section 5. Three appendices detail systematic tests of model integration accuracy (Appendix A), stability (Appendix B), and fit robustness (Appendix C). Two further appendices discuss degeneracy/biases when models fit poorly (Appendix D) and when fitting inclined thick disks (Appendix E). Lastly, Appendix F details the GalactICS method used to build galaxy models. Future papers will provide fits to a larger sample of SAMI galaxies.

2. Data

We choose a well-resolved, massive SAMI spiral galaxy (G79635), with $M_\star \approx 10^{10.4} M_\odot$ (Taylor et al. 2011) from broadband spectral energy distribution fits. Stellar kinematics are derived using single-Gaussian, two-moment pPXF

(Cappellari & Emsellem 2004) fits to the data from the blue and red arms combined, degrading the red arm (FWHM = 1.696 Å, covering the redder half of the SDSS r band) to match the blue arm's spectral resolution (2.717 Å, covering the SDSS g band); see van de Sande et al. (2017) and Fogarty et al. (2015) for details. We create "*SAMIgr*" flux maps by collapsing the spectral cube and masking emission and sky lines as in van de Sande et al. (2017). Flux uncertainties include approximate covariances (Sharp et al. 2015) added in quadrature to the shot/read noise, along with a flat systematic uncertainty corresponding to 10% (2.1%) of the faintest (peak) surface brightness. The dispersion maps exclude outliers from the best-fit radial profile. The PSF is a Moffat (1969) ellipse with 1″83 FWHM, derived via a ProFit (Robotham et al. 2017) fit to the reference star's flux map (obtained from its spectral cube exactly as for the galaxy).

g- and r-band images are from the VST-KiDS survey (de Jong et al. 2013, 2015), which covers GAMA (Driver et al. 2011) and SAMI survey regions. Uncertainties are estimated from the effective gain and local sky brightness. PSFs are Moffat ellipses with 1″16 (g) and 0″54 (r) FWHM, derived from a simultaneous ProFit fit to 39 nearby point sources. G79635 also has an H I spectrum from the ALFALFA (Haynes et al. 2011) α.70 data release.[15]

3. Methods

Our method solves a nonlinear optimization problem using a parametric galaxy model, constrained by 2D kinematic moment maps or derived quantities thereof. First, a model phase space DF must be generated (Section 3.1); second, this DF must be integrated efficiently and accurately (Section 3.2); third, the integrated DF must be projected onto a datacube (position–position–velocity) and then onto kinematic maps (Section 3.3). Finally, the optimization and sampling procedure is described in Section 3.4. Our implementation, dubbed MagRite, is based on C/C++ libraries with an R (R Core Team 2016) interface for fitting.

3.1. Galaxy Models

The models are generated using an updated version, 3.0, of the GalactICS (Kuijken & Dubinski 1995; Widrow et al. 2008) galaxy initial conditions code, to be detailed in a future paper (J. J. Dubinski et al. 2017, in preparation). GalactICS has previously been used to model the surface brightness profiles and rotation curves of local group galaxies (Widrow & Dubinski 2005; Widrow et al. 2008) and NGC 6503 (Puglielli et al. 2010), but not 2D images/kinematic maps. The core functions of the updated code are as described in Widrow et al. (2008). Key differences include the adoption of a logarithmic grid (previously linear), and the use of GNU Scientific Library (Galassi 2009) splines to create smooth differentiable functions for tabulated DFs and multipole expansion coefficients for the potential, both of which allow for more accurate function and derivative/integral evaluations using fewer grid elements than in earlier versions.

GalactICS generates equilibrium DFs with three components for galaxies:

1. An exponential stellar disk with mass $M_{d,in}$, scale radius R_d, and scale height z_d, where $\rho \propto \exp(-R/R_d)$ sech^2(z/z$_d$).
2. A (deprojected) Sérsic profile stellar bulge with scale velocity v_b and effective radius R_b, where $\rho(r) \propto (r/R_b)^{-p} \exp(-b_n(r/R_e)^{1/n_s})$, $p = 1 - 0.6097/n_s + 0.05563/n_s^2$ (Prugniel & Simien 1997), and b_n scales such that R_b is the projected half-light radius (Graham & Driver 2005).
3. A generalized Navarro et al. (1997, hereafter NFW) dark-matter halo with scale velocity v_h, scale radius r_h, where $\rho \propto [(r/r_h)^{\alpha}(1 + r/r_h)^{(\beta-\alpha)}]^{-1}$, and $\alpha = 1$, $\beta = 3$ for a "pure" NFW profile.

The minimal set of six free parameters includes a size and mass/scale velocity per component: R_b and v_b (bulge); R_d and $M_{d,in}$ (disk); r_h and v_h (halo). Four parameters control the density profiles: n_s (bulge), z_d (disk scale height), and α, β (halo); we fix $\beta = 3$ but leave the others free. We fit the disk radial central (cylindrical) radial velocity dispersion σ_{R0}, the square of which then declines exponentially with R_d. Finally, we fix the streaming fractions $f_{s,b} = f_{s,h} = 0.5$ (fractions of particles with a positive z-axis angular momentum L_z), giving non-rotating bulges and halos.

Any component c can be truncated at a radius $r_{t,c}$ with scale length $dr_{t,c}$, such that $\rho_{\mathrm{trunc}}(r) = \rho_{\mathrm{nominal}}(r)$ $[1 + \exp((r - r_{t,c})/dr_{t,c})]^{-1}$. Truncation is only strictly necessary for the halo because the NFW profile has a divergent total mass. Nonetheless, we fit the disk $r_{t,d}$ and $dr_{t,d}$ (see Section 4 for the implications of this choice), but fix the bulge $r_{t,b} = 10R_b$ ($dr_{t,b} = R_b$) and halo $r_{t,h} = 50r_{s,h}$ ($dr_{t,h} = 7.5r_{s,h}$). This adds an additional two free parameters to the previous nine.

GalactICS derives a DF for each spherical component using Eddington's formula (e.g., Binney & Tremaine 2008) and iteratively computes corrections to an analytic DF describing the disk. GalactICS then finds a potential/density pair for each component which is consistent with these DFs. The final radial density profile of each component only differs slightly from its original parametrization (see Appendix F). The key differences are that spherical components (bulge and halo) are flattened slightly in response to the presence of the disk potential and that the integrated properties of the components (e.g., disk mass M_d) can deviate slightly from their input values. Although options have recently been added to GalactICS to further correct the disk density such that the output mass and profile match the input values more closely, we omitted this step pending further testing of these new features.

Despite these caveats, Appendix B shows that under normal circumstances, GalactICS models begin in near-perfect equilibrium; perturbations on the order of a few percent result only for extreme parameter combinations. More importantly, models with large Toomre (1964) Q parameters are stable against secular evolution. Although GalactICS is not restricted to generating models with large Q—this depends mostly on the values of z_d and σ_{R0}—it can be guided to do so if necessary through priors on these input parameters or on the minimum Q value at certain radii. GalactICS also converges to a near-equilibrium DF in ∼30 s without expensive orbital integration —a major advantage over Schwarzschild/made-to-measure methods and the key requirement to permit Bayesian analyses of large samples of galaxies.
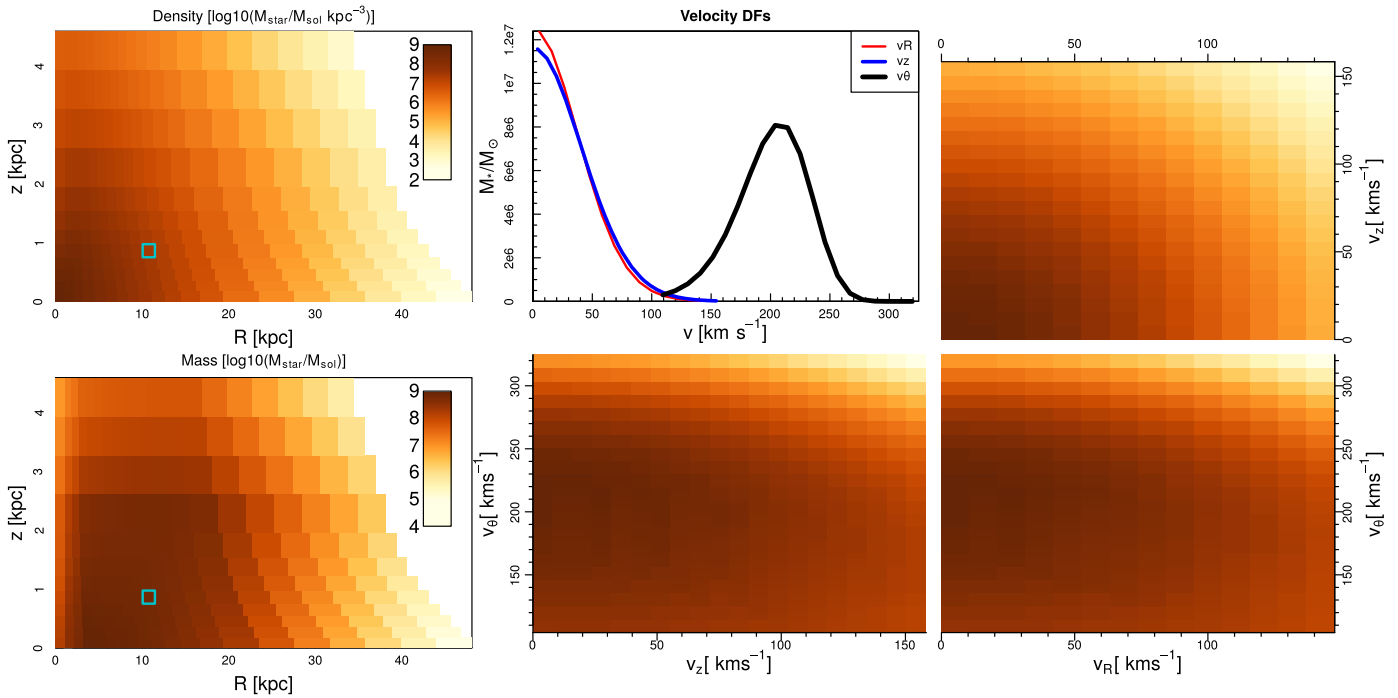
---

[15] http://egg.astro.cornell.edu/alfalfa/data/

Figure 2. Densities within selected bins of a low-resolution model grid (25 radial bins and 25 km s^{-1} velocity bins). Left panels: density and mass within the model integration grid. The scheme produces roughly equal-mass bins over most of the disk. Top: mass-weighted velocity distribution functions for the cyan-highlighted bin in the left panels. The remaining panels show pairwise (logarithmic) densities integrated over the third velocity axis, i.e., projections of the velocity "ellipsoid" (which is not perfectly ellipsoidal) at a given position in the disk.

3.2. Distribution Function Integration

The default GalactICS integration scheme samples the DF using a Monte Carlo acceptance–rejection method, which is ideal for generating unbiased, equal-mass N-body particle initial conditions. However, the rejection step is inefficient unless a suitable (i.e., strictly larger than the target distribution but only by a small margin) approximate sampling distribution is known. Unbiased sampling is not optimal for accurate integration over a uniform grid, because the fractional error is not constant but scales with density, so low-density (outer) regions have larger relative errors than the (possibly excessively) accurately sampled inner regions. Lastly, stochastic integration can induce spurious variations in the likelihood with small changes in parameter values. Evaluating the same model with a different random sequence or even a slightly different model with an identical random seed can result in spurious differences in integrated quantities and the resulting model likelihoods, depending on the number of samples. As a result, we chose to develop a faster and less stochastic grid-based integration scheme, which we will now describe in greater detail. For more detailed comparisons between these two integration schemes, see Appendix A.

We integrate the DF in its native cylindrical coordinate system and then project it rather than integrating over projected coordinates, as is usually done for 2D surface brightness profiles. For the remainder of this section, we will use the mathematical convention, where the azimuthal angle is θ, rather than the physics convention (ϕ). The disk DF is parametrized as $f(R, z, v_R, v_\theta, v_z)$. It is independent of θ and symmetric over all axes except v_θ. The DFs of the spherical components (bulge and halo) are internally functions of energy and L_z, but are reparametrized as $f(R, z, v)$ for convenience. Integrating the model over a cylindrical grid allows for some efficient

optimizations, whereas integrating down the line of sight requires repeated unique transformations at each projected position. Rotating and projecting cylindrical grid elements onto the sky plane does present a challenge. Doing so exactly requires computing the fraction of the volume of a 3D tilted ring (with a fixed height) projected within each spaxel. However, this can be roughly approximated by further discretizing each annular ring into sectors, and then assigning the mass within each sector to the spaxel containing its center of mass (in projection).

The discretized disk integration grid for G79635 is shown in Figure 2 (top-left panels). The scheme is designed to create nearly equal-mass bins. The radial grid is roughly logarithmic —each bin covers a radius containing $1/N_R$ of the total mass of an ideal, thin exponential disk. The inner and outer bins are oversampled to minimize gaps at large radii and improve accuracy near the galactic center: 15% of the bins cover the inner $0.5R_d$, whereas 35% cover $R > 3R_d$. The bins are staggered radially to spread them more evenly in projection. Vertically, the grid covers $0 < z < 10z_d$, spaced to cover equal masses until switching to linear spacing near the upper limit. For each R–z element, the disk DF is integrated over all v_R, v_z, v_θ within $(\langle v_R \rangle = 0) \pm 4\sigma_R$, $(\langle v_z \rangle = 0) \pm 4\sigma_z$, and $(\langle v_\theta \rangle \approx v_{\rm circ}) \pm 8\sigma_\theta$ and discretized into equal-velocity bins. Figure 2 shows the integration grid for a single spatial bin, including major-axis 2D projections and 1D probability distribution functions (PDFs). Typically, the DF at most spatial coordinates in the disk is nearly (but not exactly) a Gaussian ellipsoid.

The bulge uses similar radial divisions, such that the inner and outer 20% of the bins contain $0.1M_{\rm bulge}$ and $0.05M_{\rm bulge}$, respectively, accounting for the steep slope in the Sérsic profile at small/large radii for large/small values of n_s, respectively. The radial grid is divided into quadrants and then subdivided
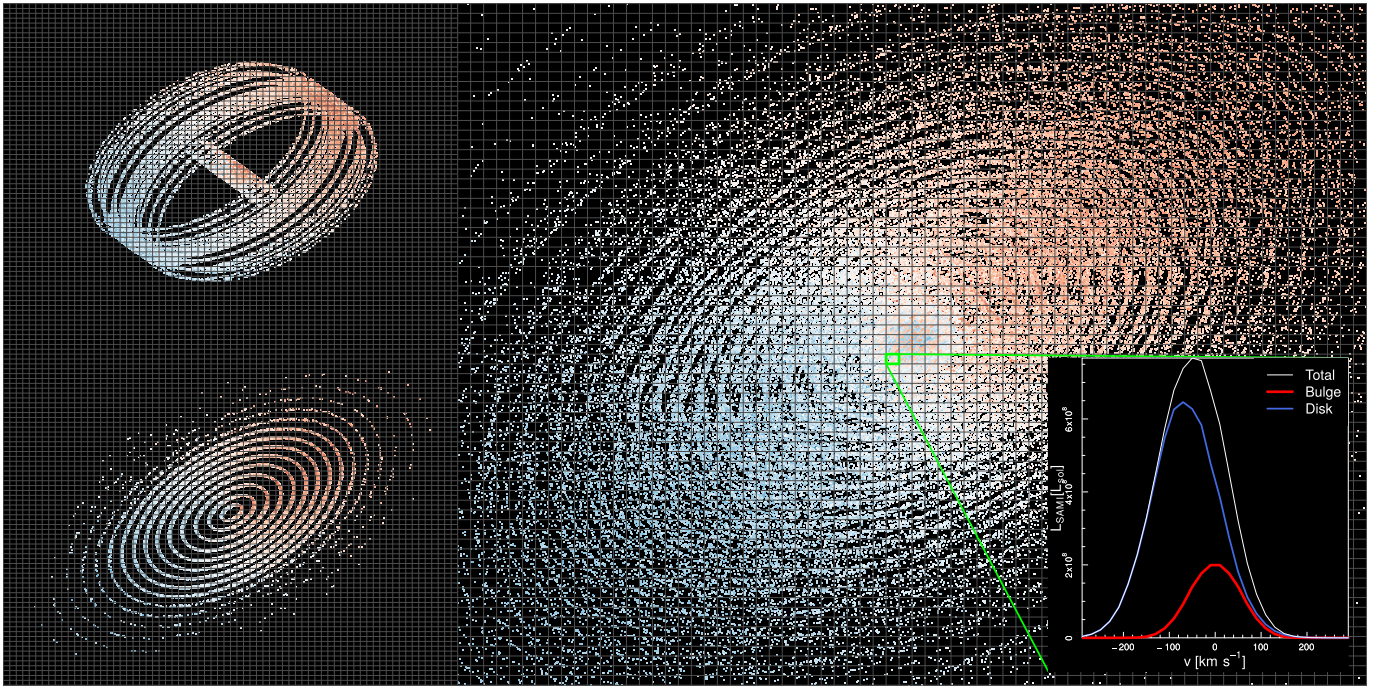
Figure 3. Visualization of the model projection and map generation schemes. Left panels: scatter plots of projected disk DF samples, color coded by line-of-sight velocity (v_{LOS}). The top-left panel shows bins at similar radii but different heights above/below the midplane, while the bottom-left panel shows a single z bin with ellipses at the same height above/below the midplane but at different radii. Right panel: scatter plot of DF samples for the disk and bulge, color coded by v_{LOS}. The inset shows v_{LOS} DFs for the highlighted spaxel (green). Radially staggered bins are distributed more evenly in projection but still create artifacts.

into linearly spaced cells along the z-axis. The bulge DF is then integrated over all $v < v_{esc}$.

3.3. Synthetic Observation Pipeline

To generate synthetic images and kinematic maps, we use an updated version of the synthetic observation pipeline described in Taranu et al. (2013) and ironically named "This Is Not A Pipeline" (TINAP). We assume an exponentially declining star formation history for the disk, SFR $\propto \exp\left(-(t - t_0)/\tau\right)$, from $t_0 = 2$ Gyr to 12.92 Gyr (the universe's age at G79635's $z = 0.04$ assuming $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\lambda = 0.7$), fitting τ^{-1} to avoid discontinuities at $\tau = 0$. The bulge is modeled as a single burst with a free formation time t_b. Both bulge and disk components have free metallicities (Z_b and Z_d, respectively). We use Maraston & Strömbäck (2011) grids to compute $M_\star/L$ in the three bands (g, r, and $SAMIgr$), assuming no stellar population gradients within components. We spawn a minimum of eight "particles" at the central R, z of each grid bin with an evenly spaced distribution from $0 < \theta < \pi/4$, beginning at a random θ (the only stochastic part of the scheme) and duplicating particles in the seven remaining octants.

The two left panels of Figure 3 show the distributions of disk particles at two fixed radii but at different heights above/below the disk midplane, color coded by v_{LOS} (top), along with particles at different radii but fixed heights above/below the disk midplane (bottom). After binning particles spatially and in v_{LOS}, every spaxel produces its own v_{LOS} PDF (right panel inset, Figure 3). These PDFs are 2D integrals of line-of-sight projections of the 3D velocity ellipsoids, and so kinematic moments are sensitive to the disk's vertical structure and anisotropy. It is worth emphasizing that Figure 2 shows a very coarse integration grid with just 25 radial bins, whereas for G79635 we use 100 bins, eliminating most discreteness effects. However, the X-shaped pattern of gaps remains even for very

fine grids. This is essentially a Moiré pattern generated by overlaying an elliptical grid onto a rectangular one. The effect is minimized but not eliminated by staggering radial bins. In practice, the patterns are small enough to be virtually invisible after PSF convolution and could be avoided entirely with more sophisticated schemes for gridding the model DF, which are under development.

For stellar kinematics, v_{LOS} cubes are convolved with the PSF and spectral line-spread function (LSF), both of which are oversampled threefold. Finally, we measure the kinematic moments in each spaxel, subtracting the LSF dispersion in quadrature for the second moment. Gaussian fits to G79635's v_{LOS} PDFs are indistinguishable from direct measurements of v_{LOS}, σ, so we use the latter. As discussed in Appendix C, this choice may not be suitable for galaxies with more massive and extended bulges, so we are investigating alternatives for future applications.

Generating synthetic kinematic maps requires nine additional free parameters—the disk inclination and position angle, offsets for x, y, v_{los}, and ages and metallicities for the bulge and disk—bringing the total to 21 free parameters.

3.4. Model Fitting

Wherever possible, initial parameter estimates and prior means are obtained from 2D ProFit fits. All priors are assumed to be normal and broad ($\sigma \approx 1$ dex). We first use a robust maximum-likelihood genetic algorithm, Covariance Matrix Adaptation—Evolutionary Strategy (CMAES; Hansen 2006). We then perform Bayesian MCMC using the Componentwise Hit-And-Run Monte Carlo (CHARM) sampler of the LaplacesDemon R package.[16] The likelihood function is the sum of the log-likelihoods from each map, assuming either a chi-

---

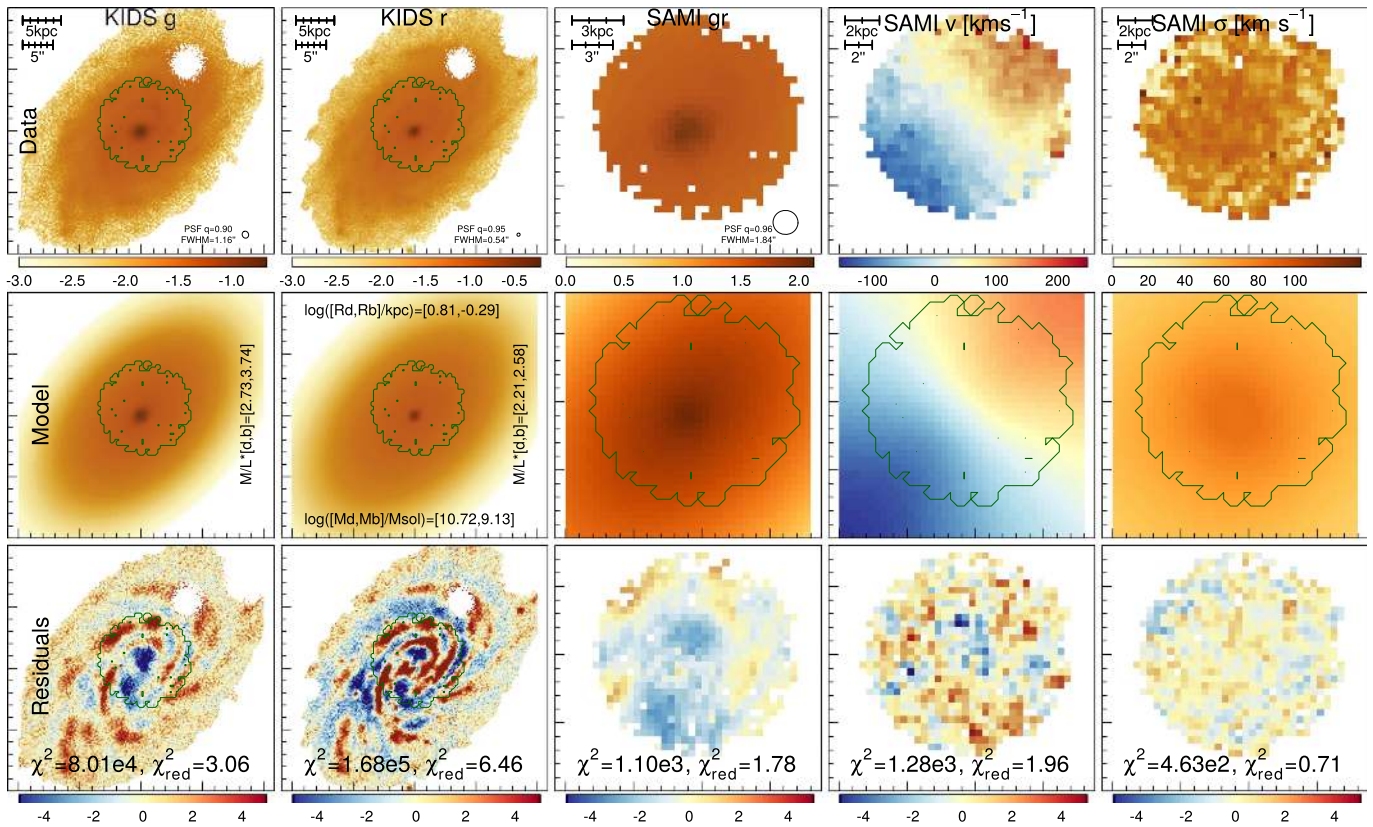[16] https://cran.r-project.org/web/packages/LaplacesDemon/

Figure 4. Best-fit G79635 model using SAMI moment 0–2 maps and KiDS $g + r$ images, along with residuals relative to pixel/spaxel uncertainties. KiDS images are 200×200 ($2''$ ticks) while SAMI images are 36×36 ($1''$ ticks).

square distribution for image residuals or a sum of normally distributed residuals for kinematic maps with less well-defined errors. Residuals are defined as $\chi_i = (\mathrm{data}_i - \mathrm{model}_i)/\mathrm{error}_i$ and $\chi^2 = \Sigma_i((\mathrm{data}_i - \mathrm{model}_i)/\mathrm{error}_i)^2$. Note that although we quote reduced χ^2 (χ^2_{red}) values, they are not used in the optimization procedure, which instead derives likelihoods from the chosen statistics' PDF directly (i.e., by calling the "dnorm" and "dchisq" functions in R).

Our CMAES code is based on the R CMAES package.[17] We have modified both CMAES and LaplacesDemon, implementing runtime limits for supercomputer queues; these versions are available on GitHub.[18,19]

4. Results

The best fit for G79635 using SAMI and KiDS $g + r$ is shown in Figure 4. The χ^2_{red} for all of the flux maps is significantly above unity. However, the largest r-band residuals clearly trace non-axisymmetric features like spiral arms and interarm gaps, and the similarity in residuals across independent data sets is encouraging, given the systematics introduced by SAMI's cubing procedure and single-star flux calibration. The r-band fit is worst simply due to its higher signal-to-noise ratio (better seeing and longer exposures than g).

Figure 5 shows 1D profiles azimuthally averaged over the best-fit ProFit disk ellipse, compared to a ProFit 2D double Sérsic r-band fit with a free bulge position angle. The dispersion map/profile is overfit and the best-fit rotation curve appears to rise

---

[17] https://cran.r-project.org/web/packages/cmaes/
[18] https://github.com/taranu/LaplacesDemon
[19] https://github.com/taranu/cmaeshpc

slightly too steeply, as can also be seen in Figure 4 (where the velocity map residuals show spatial coherence). Encouragingly, the predicted rotation curve at a fiducial radius of $(3–3.4)R_d$ (Catinella et al. 2007) is within 10% of the independent ALFALFA H I $W_{50} = (347 \pm 8)$ km s^{-1} measurement, even though the H I data were not used in the fit and the SAMI data does not appear to reach the peak of the rotation curve. The lower stellar velocity could be due to asymmetric drift, as it is not unusual for stellar disks with radial dispersion support to have $\sim$10% lower rotation speeds than gaseous disks (Ciardullo et al. 2004; Martinsson et al. 2013; Brooks et al. 2017). The peak stellar velocity is also consistent with the independent $V_{\mathrm{max}} \sin i = 165$ km s^{-1} circular speed derived by Cecil et al. (2016).

The fact that the observed mean stellar velocity lies well below the circular speed curve is due to a combination of factors. First, the mean velocity within $R_{e,\mathrm{disk}}$ is decreased due to beam smearing (compare the solid blue and dashed blue lines in Figure 5). This effect is modest beyond the peak of the rotation curve (compare the solid and dashed rotation curves), although it continues to boost the observed velocity dispersion by about 10 km s^{-1}. Note that estimates of the mean velocity and dispersion are unreliable beyond about 15 kpc, where the dispersion drops well below the 60 km s^{-1} velocity grid resolution. Also, there is some subjectivity in how 1D apertures are defined. We measure velocities and velocity dispersions using data from spaxels within 5 and 10 degrees of the major axis, respectively. The model-projected velocity without beam smearing (dashed blue curve in Figure 5) is measured within the same apertures as the PSF-convolved version (solid line) and does not take into account the fact that PSF convolution
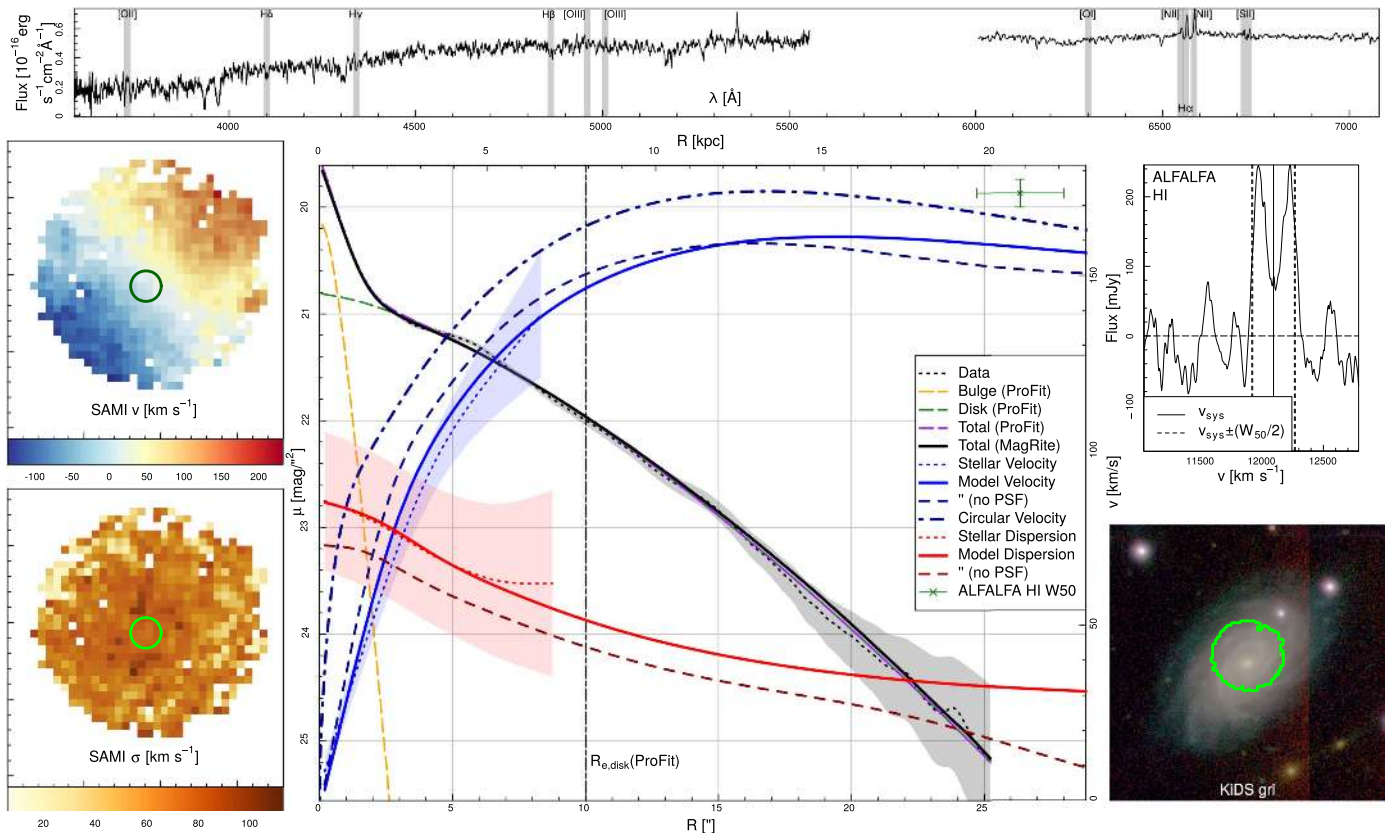
Figure 5. Data and 1D profiles for G79635. Clockwise from bottom left: the SAMI dispersion and velocity maps, with a $2''$ diameter aperture (green); the SAMI spectrum within this aperture, with emission lines excluded from the fit shaded in gray; the ALFALFA H I spectrum; RGB composite image using KiDS g-, r- and i-band images, overlaid with SAMI coverage (green). Center: azimuthally averaged, spline-fit 1D profiles (dotted lines) for the first three kinematic moments (KiDS r-band surface brightness—gray/black; velocity—blue; velocity dispersion—red), along with 1σ uncertainties (shaded). Also shown are the best-fit MagRite (thick solid lines) and ProFit 2D double Sérsic r-band fits (dotted lines), split among bulge (orange), disk (green), and total (purple). The background gradient in the KiDS images (most visible in the red i-band channel) was removed prior to fitting. Note that velocities are as observed and not inclination corrected, i.e., $v \approx v_{circ}\sin(i)$, where v_{circ} is the circular velocity and i is the disk inclination. The best-fit MagRite model $v_{circ}\sin(i)$ (dotted–dashed dark blue line) lies well above the observed rotation curve, illustrating the combined effects of beam smearing, asymmetric drift, and a thick disk.

modifies isophotes as well; however, since the 1D kinematics are measured close to the major axis, this effect is minor.

In the inner few kiloparsecs, the mean velocity is suppressed both because of the presence of a non-rotating bulge and because the disk has a finite thickness, so that a large fraction of disk stars are a significant distance away from the disk midplane. Beyond this inner region, asymmetric drift and the non-zero radial and vertical velocity dispersion of the disk continue to lower the mean velocity. Observations suggest that it is not unusual for stellar disks with radial dispersion support to have ~10% lower rotation speeds than gaseous disks (Ciardullo et al. 2004; Martinsson et al. 2013; Brooks et al. 2017).

Despite also fitting the g-band image and SAMI kinematics, the MagRite best-fit model is a better fit to the KiDS r-band image than a single-band exponential disk ProFit fit; ProFit only fits slightly better with a free n_s disk.

Table 1 lists best-fit values and uncertainties for the MagRite model parameters and several key derived quantities. Posterior distributions for selected common parameters of the MagRite and ProFit fits are shown in Figure 6. We find that direct fits to the data yield unreasonably narrow PDFs, listed as $\sigma_{obs.}$ in Table 1. To test whether MagRite is the cause of this effect, we generate and fit noisy mock maps of the best-fit model (see Appendix C for a full description of the procedure). We find no evidence for significant bias in the best-fit parameter values.

This form of "noise bias" can be significant in a low signal-to-noise image, as is the case in weak lensing studies (e.g., Bernstein & Jarvis 2002; Refregier et al. 2012). However, the parameter PDFs for fits to the mock maps are significantly broader (Table 1) than when fitting the actual data—in some cases by over two orders of magnitude—and ProFit exhibits similar behavior.

As Figure 4 shows, an axisymmetric disk is not a good fit to the flux maps and cannot reproduce the spiral arm structure evident in the KiDS images (especially in r). In general, models that fit data poorly underestimate uncertainties significantly, although the degree to which this occurs does depend on the model and fit statistic. This result is not immediately obvious, and we discuss it further in Appendix D. Our solution of fitting mock images to obtain more realistic parameter uncertainties is necessary but likely insufficient. That is, the σ_{mock} in Table 1 should be interpreted as a lower bound on the uncertainty on each parameter in the highly idealized scenario that the galaxy is perfectly described by the model. There is no obvious prescription for estimating or adjusting parameter uncertainties for models that do not fit the data well.

Table 1 also lists Δ, the difference between the ProFit and MagRite best-fit values for common parameters (whether derived or fit directly). This can be considered as an estimate of systematic uncertainties from using two different (but still similar) modeling methods. In all cases, $|\Delta|$ is larger than

Table 1
Best-fit G79635 MagRite Parameters

Name	Unit	log[a]	Mean	$\sigma_{obs.}$	σ_{mock}	Δ[b]
Fitted Parameters						
$M_{d,in}$	M_{sim}[c]	Y	1.604	8.76e−5	1.33e−3	...
R_d	kpc	Y	8.141e−1	2.62e−5	9.85e−4	...
z_d	kpc	Y	2.224e−1	1.83e−4	8.08e−3	...
$r_{t,d}$	kpc	Y	1.146	2.07e−4	1.67e−3	...
$d\check{r}_{t,d}$	kpc	Y	7.065e−1	3.56e−4	5.26e−3	...
σ_{R0}	v_{sim}[d]	Y	−1.281e−1	2.93e−4	2.44e−2	...
v_b	$\sqrt{v_{sim}}$	N	1.020e−2	1.80e−4	1.92e−3	...
R_b	kpc	N	−2.899e−1	1.01e−4	7.35e−3	−3.62e−2
n_s	N/A	Y	−9.970e−2	9.79e−4	2.45e−2	1.82e−1
v_h	$\sqrt{v_{sim}}$	Y	2.693e−1	6.41e−5	1.82e−3	...
r_h	kpc	Y	8.051e−1	3.17e−4	1.47e−2	...
α	N/A	N	9.845e−1	1.57e−4	2.31e−2	...
τ^{-1}	Gyr^{-1}	N	4.851e−1	1.36e−3	6.62e−3	...
t_b	Gyr	Y	8.450e−1	1.19e−3	8.55e−3	...
Z_d	$Z/Z_\odot$	Y	−5.170e−1	1.75e−3	7.19e−3	...
Z_b	$Z/Z_\odot$	N	3.000e−1	8.17e−5	8.83e−3	...
P.A.	rad.	N	8.059e−1	2.72e−4	2.08e−3	−3.26e−3
$\sin(i)$	rad.	N	8.160e−1	1.55e−4	1.65e−3	2.09e−2
X_{off}	kpc	N	3.732e−2	4.32e−4	5.81e−3	5.11e−2
Y_{off}	kpc	N	1.513e−2	9.68e−4	5.89e−3	5.91e−2
$V_{z,off}$	km s^{-1}	Y	−1.547e−1	3.47e−1	3.55e−1	...
Derived Parameters						
M_d[e]	$M_\odot$	N	10.72	1.43e−4	1.38e−3	...
M_b	$M_\odot$	N	9.126	6.93e−4	1.07e−2	...
$(M/L_g)_d$	$(M/L_g)_\odot$	N	2.726	1.47e−3	1.18e−2	...
$(M/L_r)_d$	$(M/L_r)_\odot$	Y	2.213	8.60e−4	7.42e−3	...
$(M/L_g)_b$	$(M/L_g)_\odot$	N	3.745	1.18e−2	8.79e−2	...
$(M/L_r)_b$	$(M/L_r)_\odot$	N	2.581	7.44e−3	5.42e−2	...
L_d	$L_{\odot,r}$	Y	10.38	1.54e−4	8.76e−4	4.46e−3
L_b	$L_{\odot,r}$	Y	8.715	9.91e−4	7.58e−3	2.99e−2
$R_{e,d}$[d]	kpc	Y	0.9365	8.82e−5	1.10e−3	7.10e−3
M_h	$M_\odot$	Y	11.81	3.03e−4	3.33e−2	...

Notes. Fitted parameters are listed as fit internally by MagRite, whereas derived parameters are measured during or after model generation.
[a] Values listed as $\log 10$(value unit^{-1}).
[b] MagRite value less ProFit value (where applicable), where the ProFit model has a thin Sérsic disk sharing its position angle with the bulge.
[c] $M_{sim} = 2.3245e9\ M_\odot$.
[d] $v_{sim} = 100$ km s^{-1}.
[e] Directly measured model half-light radius, accounting for truncation.

σ_{mock}—sometimes by more than an order of magnitude. This suggests that systematic uncertainties dominate over statistical uncertainties. Unfortunately, we are unaware of any robust methods for incorporating systematic uncertainties into our likelihood functions, so the only obvious solution to this issue remains increasing the model's flexibility until it can reproduce the data.

Our testing demonstrates that MagRite will recover input parameters correctly from idealized mock data, but this does not guarantee realistic parameter values when fitting real galaxies. For example, the MagRite model has an unrealistically large disk scale height, $z_d = 1.67$ kpc, and a small truncation radius, $r_{t,d} = 14.0$ kpc, compared to the scale length

$R_d = 6.52$ kpc. These values seem to compensate for features in the data not otherwise described by the model. G79635's disk appears steeper than exponential, and the best-fit ProFit disk $n_s \approx 0.8$ (Figure 5). The small truncation radius steepens the MagRite surface brightness profile at large radii, whereas the large scale height lowers the surface brightness along the minor axis from the galaxy center—precisely where there are two underdense interarm gaps. A model with azimuthal variations and a more general Sérsic or broken-exponential disk profile might prefer a thinner, non-truncated disk. Having said that, Muñoz-Mateos et al. (2013) fit broken-exponential profiles to Spitzer 3.6 μm imaging of nearly face-on disks and found a typical break radius at 2.3 ± 0.9 inner scale lengths, so the truncation radius is not unreasonable for a Type II (truncated, as per Freeman 1970) disk.

The disk mass is also considerably higher than the total stellar mass estimated by Taylor et al. (2011) from fits to photometry alone. G79635 has a rather large estimate H I mass of $10^{10.22}\ M_\odot$, so it is possible that our larger disk mass is compensating for the contribution of the gas disk to the rotation curve. This could also be the cause of the slight underprediction of the rotation curve at large radii, if it is not due to asymmetric drift. In practice, a more flexible and better fitting stellar mass model would likely allow the halo parameters to vary more to compensate for such inconsistencies.

The best-fit disk metallicity is quite low ($\log 10(Z_d/Z_\odot) < -0.5$) for such a massive disk, whereas the bulge metallicity reaches the ceiling of the Maraston & Strömbäck (2011) model grids ($\log 10(Z_b/Z_\odot) = 0.3$). By contrast, the disk is fairly old, with a short $\tau = 2.06$ Gyr, while the bulge has a moderate age of 5.92 Gyr. The observed galaxy colors cannot be reproduced by such a relatively simple model; in particular, the galaxy center is redder than the model, and the outskirts are significantly bluer, both by about 0.2 in $g - r$ and with a fairly sharp transition rather than a smooth gradient. Additional model complexity (especially dust reddening and stellar population gradients) is necessary to fully reproduce galaxy colors, and full spectral modeling would be ideal. However, it is worth noting that systematic differences in the inferred stellar masses of GAMA galaxies (including the SAMI sample) can be as large as 0.2 dex depending largely on the treatment of star formation histories and dust (Wright et al. 2017), even neglecting possible variations in the initial mass function. There are also significant differences among stellar population models, stellar spectral libraries, and isochrones, which preclude making accurate estimates of stellar mass-to-light ratios even given a star formation history, and it is unclear how one might estimate the magnitude of such effects for a given galaxy.

One potential general solution to limit parameter bias is to introduce stricter priors on model parameters based on external data. Disk scale height distributions can be constrained from independent observations of edge-on disks, and dust extinction can be estimated from Balmer line decrements. Ultimately, the best solution is to improve the model itself, which would permit the quantification of such biases. Such improvements are planned for MagRite but are not necessary to implement the method itself. For the moment, we advise caution when interpreting uncertainties from models that do not adequately reproduce known or clearly visible features in the data. This particular galaxy is well-resolved compared to average SAMI galaxies (although not uniquely so); these issues are less
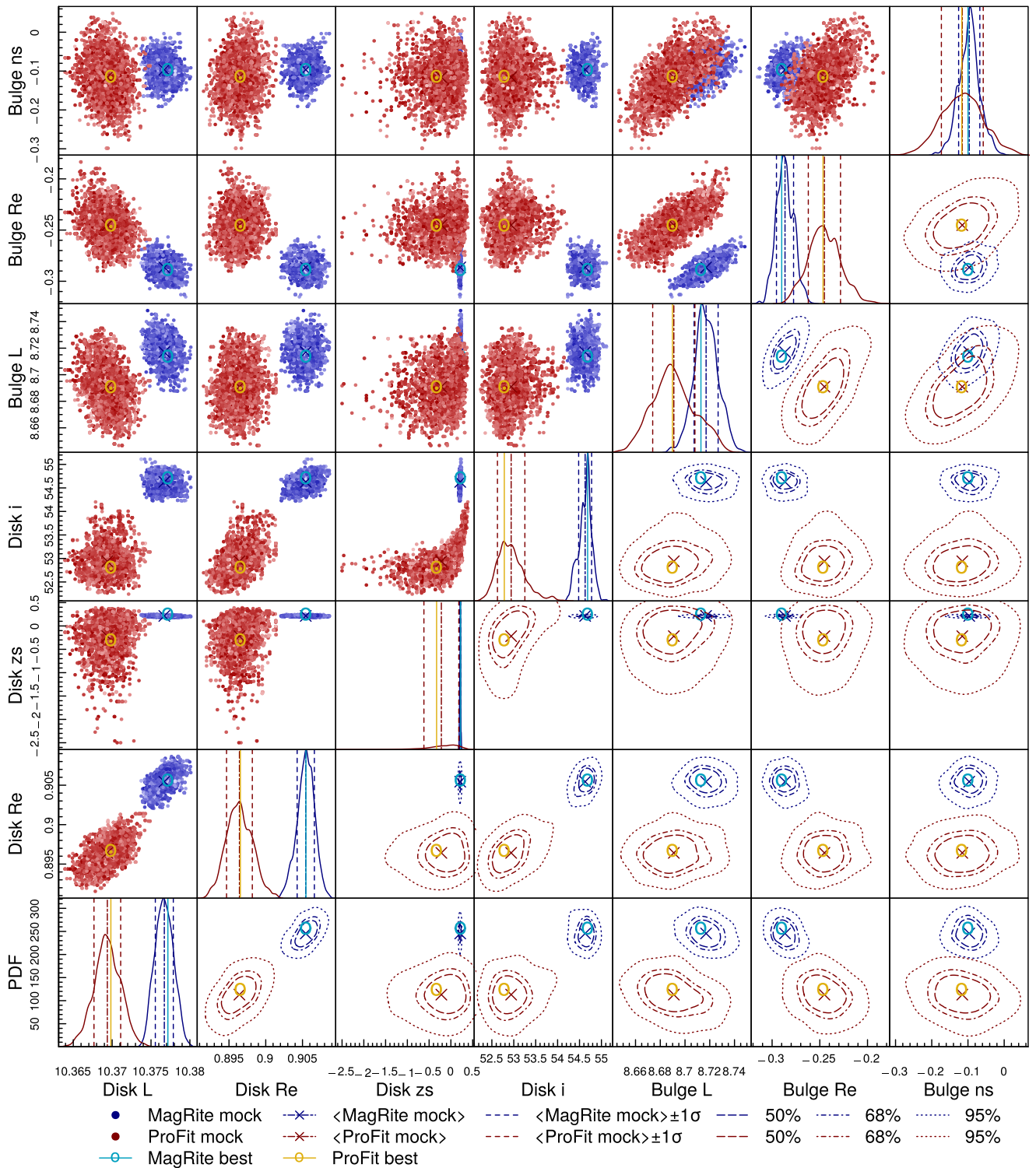
Figure 6. Triangle plot showing joint posterior parameter distributions ($L \equiv \log 10(L_r/L_\odot)$, $Re \equiv \log 10(R_e/\mathrm{kpc})$, $zs \equiv \log 10(z_d/kpc)$, $n \equiv \log 10(n_s)$, and i is the inclination in degrees) for ProFit (blue) and MagRite (red), where the ProFit disk is a Sérsic profile and shares its position angle with the bulge. The upper-left half shows the scatter plots of accepted samples, while the bottom-right half shows the 1D and smoothed 2D probability contours. Accepted samples are colored by probability on an arbitrary scale, such that more probable points have darker and more saturated colors. Plots along the diagonals show the PDFs of accepted samples for the variable on the x-axis; to avoid crowding, the y-axis ticks and labels are omitted for the three interior histograms. All posterior distributions are for fits to mock data; the best-fit parameters used to generate the mock data are plotted as circles. Note that because the ProFit thick disk fits have an unrealistically large disk scale height, the ProFit mock uses the best-fit thin disk fit parameters with $z_d = 0.1R_{e,d}/1.67835$ (equivalent to $0.1R_d$ for an exponential disk); this illustrates the limited constraints on disk thickness from photometry alone.

pronounced when fitting lower-quality data. However, as the recent public release of Subaru Hyper Suprime-Cam data (Aihara et al. 2017) shows, high-quality, deep ground-based imaging is rapidly becoming available for large galaxy surveys —even in the southern sky (Keller et al. 2007)—and so this issue cannot be ignored for much longer.

5. Conclusions

We outlined a method for kinematic bulge–disk decomposition using self-consistent, DF-based dynamical models. The method can be used to model any combination of data, including deep optimal images and 1D/2D kinematic constraints. Our GalactICS-based implementation (MagRite) is efficient enough (∼1–2 minutes per model on modern CPUs) to fit deep KiDS images and SAMI kinematic maps (Figure 5), exactly as conceptualized in Figure 1.

We fit the well-resolved SAMI/GAMA galaxy G79635, showing that the best-fit parameters and posteriors are largely consistent with ProFit 2D decompositions and with the independent H I W_{50} constraint. This suggests that MagRite can extrapolate reasonable rotation curves even without IFS data reaching the peak of a galaxy's rotation curve—a crucial requirement for accurate stellar mass and angular momentum estimates. In the provided appendices, we demonstrate that MagRite can fit synthetic model data with minimal biases. However, we caution that fits to real data are not immune to biases, particularly in the presence of significant non-axisymmetric features. Furthermore, we showed that poorly fitting models can seriously underestimate parameter uncertainties by yielding artificially narrow posterior PDFs. This can be mitigated, but not corrected, by estimating uncertainties from mock observations of the best-fit model.

Our example galaxy was selected as a well-resolved, fairly passive spiral galaxy with an H I detection, but there are many more SAMI galaxies with similar quality data. The KiDS images are good enough to constrain the bulge fraction in G79635 to at most a few percent and clearly show deviations from our idealized models, which assume axisymmetric, exponential disks and simple star formation histories for each component. We therefore demonstrate that data quality is not the main impediment to improved physical modeling of galaxies, but the models themselves. G79635's azimuthally averaged disk profile can be reproduced with a combination of an unusually thick and smoothly truncated exponential disk, but would be better fit with a Sérsic or non-parametric profile disk including perturbations from spiral arms.

One shortcoming of the method using kinematic moment maps is that these must be derived independently; nonetheless, the method can be generalized to fit spectral datacubes directly (e.g., Tabor et al. 2017), and we plan to implement this functionality within MagRite. Additional model features like spiral arms, dust attenuation/scattering (Pastrav et al. 2013), and more flexible/non-parametric density profiles are longer-term ambitions. MagRite is under active development and will be released in the near future, alongside early results from a larger SAMI sample. Parties interested in testing the code or contributing to future development are encouraged to contact the authors.

Appendix A
Integration Scheme Comparison

To test the accuracy and speed of the DF integration scheme described in Sections 3.2 and 3.3, we generate synthetic maps using our method and also from a high-resolution GalactICS model with 20M/0.4M disk/bulge particles, respectively. Figure 7 shows the maps and residuals generated using the best-fit model parameters for G79635 with both of these integration schemes. Despite the large number of particles, the MC (N-body) maps are still shot noise limited near the outskirts of the disk. Furthermore, sampling this many particles takes nearly 10 minutes on a modern test machine (Intel i5-4690 at 3.50GHz). Each accepted particle requires just over three proposals on average, meaning that nearly 70% of the computing time is effectively wasted evaluating rejected proposals. In contrast, the grid-based integration method takes under a minute and generates smoother, noise-free maps that are virtually indistinguishable from the unbiased N-body maps near the well-sampled galaxy center. This is accomplished mainly by making fewer calls to the expensive DF evaluation methods, effectively spawning dozens to hundreds of particles per DF sample.

As discussed in Section 3.3, our grid-based integration scheme is not entirely ideal. Placing evenly distributed samples at the center of each bin is computationally efficient but requires a random angular offset in θ to generate smooth maps —otherwise, the model images would have bright "spokes" at the sampled angles θ and artificial gaps between them. Similarly, binning evenly spaced ellipses onto a rectangular grid creates Moiré-like artifacts, apparent as an X-shaped residual in Figure 7. These issues could be resolved by distributing DF samples over the projected areas of elliptical rings rather than as points binned onto a rectangular grid. In principle, this would also need to be done in 3D, i.e., by generating rings corresponding to the top of one bin and the bottom of the next. These would be fairly inexpensive calculations compared to the other steps in model integration, but are somewhat complex and would be unlikely to change the PSF-convolved model maps significantly, so they are left to future revisions of MagRite.

One measurable impact of the integration scheme is the change in model likelihoods through stochasticity. To test this, we generate a series of maps for the best-fit model, varying only the random seed used to generate the angular offsets in θ. For the r-band KiDS image, the χ^2 value varies by about 40 from different seeds. This is an insignificant difference for a well-fitting model, but highly significant for one with a large χ_{red}^2, as discussed in Appendix D. To minimize the impact of this issue, we keep the random seed fixed for all fits.
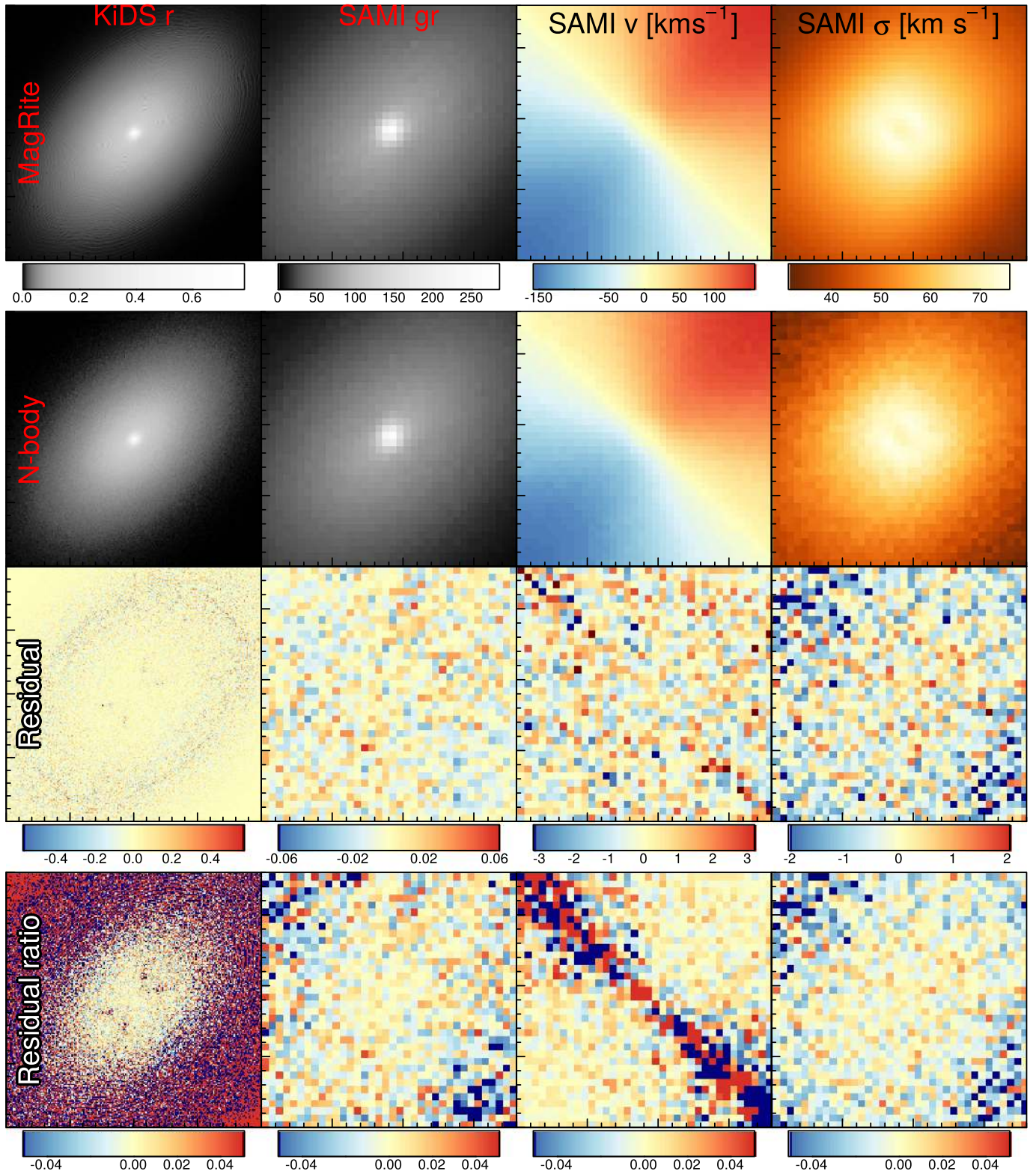
Figure 7. Comparison between maps from grid-based and Monte Carlo integration using 20M/0.4M disk/bulge particles, respectively. The residuals are shown on an absolute scale (MagRite—*N*-body) and as a ratio relative to the smoother MagRite map. The large relative velocity residuals along the minor axis are due to the small absolute value of the velocity; absolute differences are small.

Appendix B
Model Stability Test

To test the long-term stability of the model, we generated *N*-body initial conditions with GalactICS, sampling the disk/ bulge/halo with 5M/0.1M/5M particles and using softening lengths of 50/50/150 pc, respectively. We ran the model for 1 Gyr with PARTREE (Dubinski 1996), using a fixed 0.196 Myr time step and an opening angle of 0.8. Figure 8
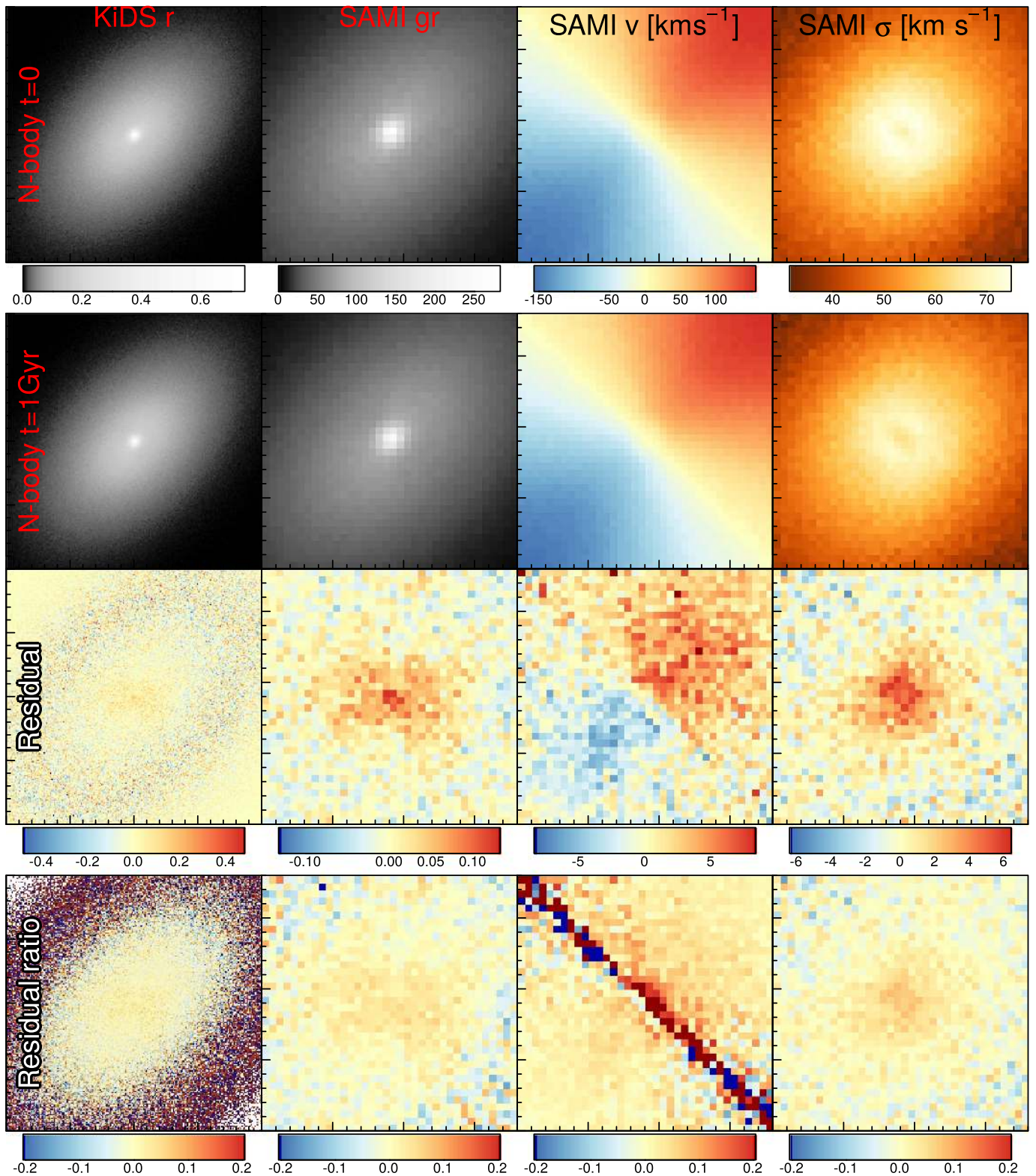
Figure 8. Comparison between maps generated from *N*-body initial conditions and after 1 Gyr of evolution. Pixels outside of the limits of the color bar are assigned the darkest (most saturated) colors. The residuals show some evolution in the galaxy structure due to the thick, truncated disk being slightly out of equilibrium.

shows the resulting maps from the evolved galaxy compared to the initial conditions. There is evidence of relaxation of the system, with the evolved model having lower central density and velocity dispersion and a shallower rotation curve.

The evolution of this model is not representative of a typical GalactICS model. As discussed in Section 4, the best-fit model has an unusually large disk scale height and small truncation radius. The initial virial ratio $q = -2T/W$, where T and W are the total kinetic and potential energies, respectively, is 1.00275. While the deviation from unity is not large, the total energy of the system is dominated by the dark-matter halo (with nearly 90% of the mass), so the stellar component is

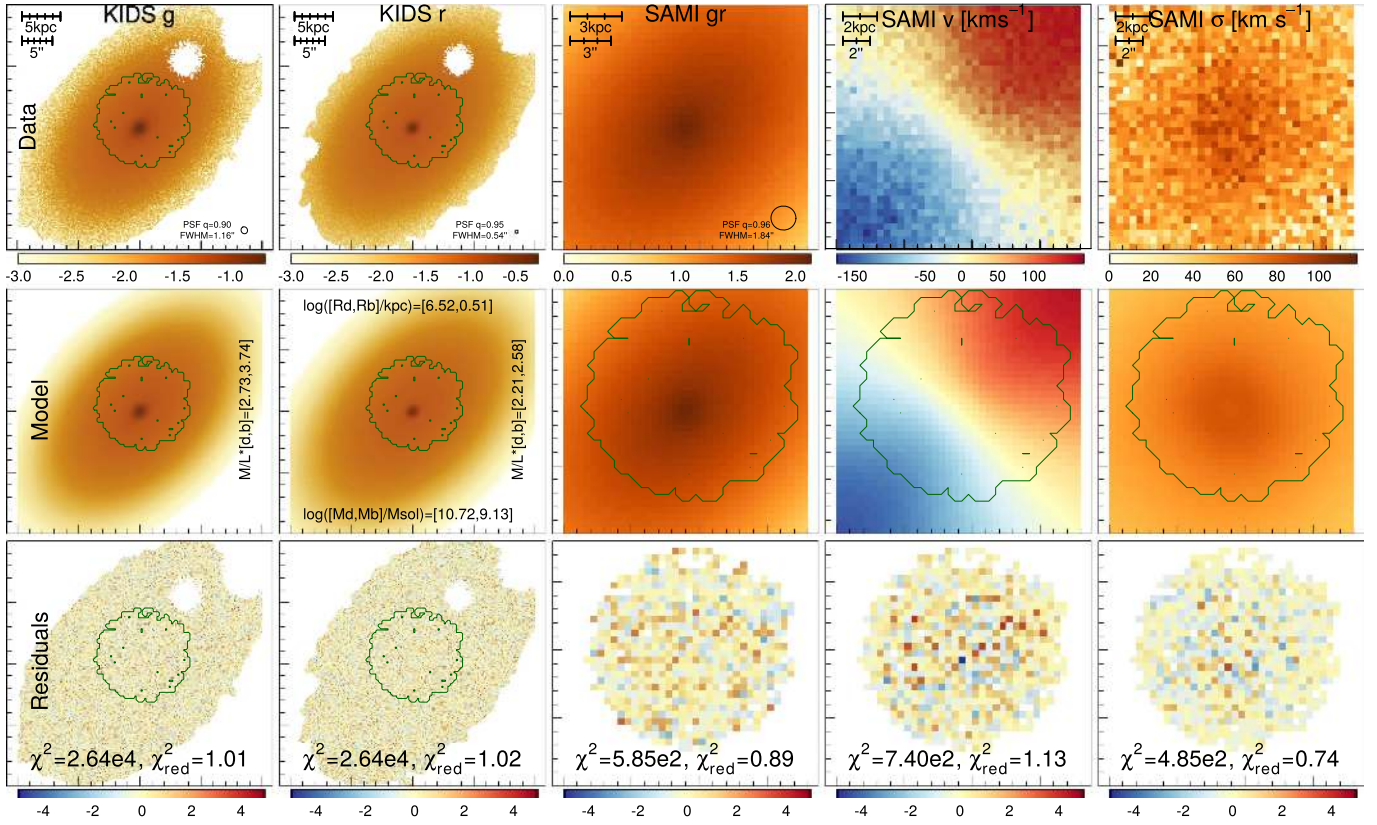**Figure 9.** Mock G79635 maps with realistic shot noise. The kinematic maps are derived from Gaussian fits to noisy v_{los} DFs and therefore do not necessarily follow Poisson or (approximately) normal statistics.

likely out of equilibrium by a few percent. Accordingly, the virial ratio drops below unity by a similar factor and shows damped oscillations before reaching a new equilibrium.

We adjust the model to have a thinner disk ($z_d = 1 kpc$) and much larger truncation radius ($r_{t,d} = 8 r_d$) with the same disk mass. This model shows virtually no evolution outside of the inner 200 pc, where there is a modest depression in the central density and velocity dispersion. We conclude that while the disk truncation parameters and scale height can in principle mimic a non-exponential ($n_s < 1$) disk, adjusting them beyond GalactICS limits can produce unstable models and should be avoided. This could be accomplished without running simulations simply by placing stronger priors on model parameters or on output diagnostics like the virial ratio, but further testing is needed to determine guidelines for these limits.

Appendix C
Model Fitting Test

We test the code's ability to recover model parameters by fitting synthetic maps generated by MagRite. We assume shot noise-dominated errors for the flux maps, given a gain and mean sky brightness in counts per pixel. The higher-order SAMI moment maps require some simplifying assumptions. Kinematic constraints originate mainly from stellar absorption lines, which only cover a small fraction of typical spectra. We parametrize this effect with a simple "kinematic gain" ratio $g_{k,\mathrm{eff}}$, which is roughly the ratio of the sum of the equivalent widths of all absorption lines to the full wavelength range. We generate a noisy flux-weighted v_{LOS} DF for each spaxel, where

the counts in each bin are multiplied by $g_{k,\mathrm{eff}}$, and fit a Gaussian to extract the mean velocity and dispersion. Finally, we reuse the existing masks and also the velocity and dispersion error maps for consistency with the original fits. We also adjust $g_{k,\mathrm{eff}}$ until the noise in the velocity/dispersion maps is roughly consistent with the original errors.

Figure 9 shows synthetic noisy maps for G79635 using $g_{k,\mathrm{eff}} = 0.02$. As expected, the flux maps are completely consistent with (nearly) normal shot noise and have $\chi^2_{\mathrm{red}} \approx 1$. However, the noise in the v_{LOS} and σ maps is not entirely identical to that from the original SAMI maps, being slightly over- and underestimated, respectively. This is not surprising, given the nonlinear nature of kinematic fits, but the mock kinematics still appear consistent with random noise without any obvious systematic bias and are usable as data for a mock fit. The possibly conservative choice of $g_{k,\mathrm{eff}} = 0.02$ is motivated by the need to keep this noise consistent with the SAMI error maps; larger values lead to excessively smooth kinematic maps. When fitting, we continue to use directly measured velocity moments in the model rather than the Gaussian fits used for both the mock and real data. This is because fitting v_{LOS} requires some estimate of the uncertainty of the v_{LOS} DF in each spaxel, and it is not clear what this uncertainty should be in a noiseless model. Since G79635 has a very modest bulge, there is little difference between the two measurements; however, this may not be the case in galaxies with more massive and extended bulges, where this issue would be worth revisiting.
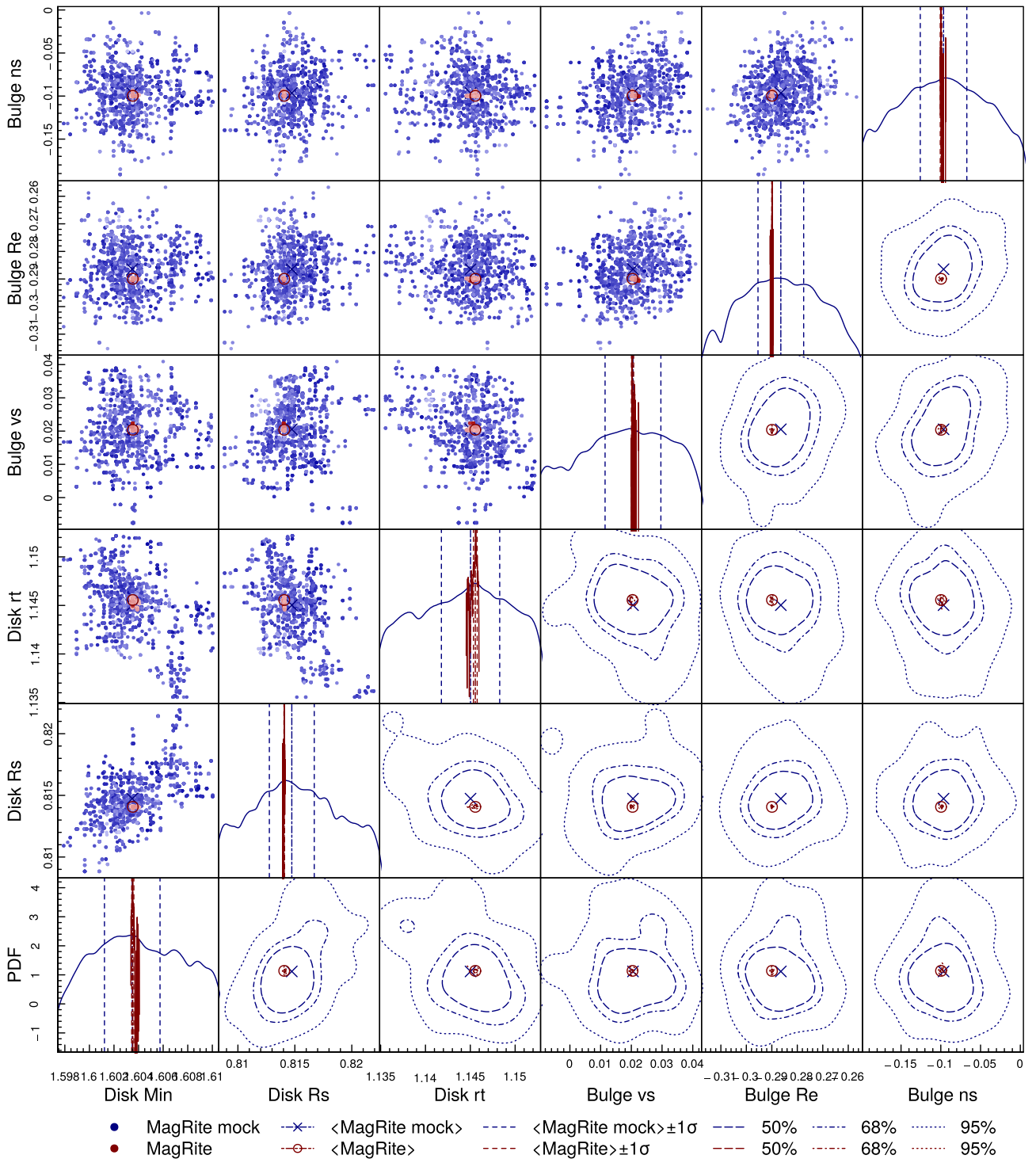
Figure 10. Triangle plot showing joint posterior distributions for model input parameters (Disk Min ≡ $\log 10(M_{d,in}/M_{sim})$, Disk Rs ≡ $\log 10(R_d/\mathrm{kpc})$, Disk rt ≡ $\log 10(r_{t,d}/\mathrm{kpc})$, Bulge vs. ≡ $\log 10(v_b/v_{sim})$, Bulge Re ≡ $\log 10(R_b/\mathrm{kpc})$, ns ≡ $\log 10(n_s)$). Points are color coded by log probability as described in Figure 6. PDFs are shown for fits to the mock data and the observed data. Since the latter are exceptionally narrow, univariate PDFs along the diagonal are shown on a logarithmic scale spanning six orders of magnitude.

Figure 10 shows the posterior distributions for a MagRite fit to the mock data shown in Figure 9. Reassuringly, the best-fit parameters are close to the inputs, with small deviations well within the 1σ uncertainties. Small offsets are expected given the noise and the difference between the directly measured and fitted kinematics, but there are no significant systematic biases.

This is not to say that fits to real data will be devoid of biases—this simply verifies that MagRite recovers unbiased parameters when the data are perfectly described by the model.

One possible concern is that there is still some noise in the model probabilities themselves—when colored by log probability, the points do not show smooth gradients in posterior

probability from the maximum-likelihood solution. This is especially true for the fits to the observed data, which is unfortunately not clearly visible in Figure 10 because the posteriors are so narrow. This is likely due to the issues with model integration outlined in Appendix A. The main practical effect of a slightly stochastic model likelihood is that convergence to the best-fit solution can be slow, as the solution wanders between entirely artificial local maxima. This problem is exacerbated when the best-fit model is a poor fit, since even extremely small changes in input parameters can cause spurious changes in the model likelihood. Fortunately, since the actual changes in parameter values tend to be small, the impact on the posteriors from mock fits is minimal.

Since poorly fitting models tend to vastly underpredict parameter uncertainties (for reasons detailed in Appendix D below), we suggest that uncertainties from mock fits should be considered lower bounds on the "true" parameter uncertainties. Again, this does not guarantee that there are no biases in the best-fit parameters themselves, so caution must be taken in defining priors and interpreting best-fit values in such cases. Similarly, since the mock data are idealized, they should only be interpreted as lower limits until further testing is done to quantify the impact of deviations from the assumed model parametrizations, which are necessarily present in all galaxies.

Appendix D
Parameter Uncertainties from Poorly Fitting Models

It is not necessarily intuitive that poorly fitting models can or should yield systematically smaller uncertainties than good models. Nonetheless, it is a natural consequence of the shape of some common statistical distributions. The difference in the log-likelihood (ΔLL) between an $n\sigma$ and $(n + 1)\sigma$ deviation from a univariate normal distribution is $-(n + 0.5)$. The difference between 10σ and 11σ deviations is then much more significant than that between 2σ and 3σ deviations, simply because the log of the PDF of a normal distribution declines as $x^{-2}/2\sigma$.

For a more practical example, the KiDS r-band image for G79635 has 25,961 usable data points (unmasked pixels). Using the chi-square distribution with 25,961 degrees of freedom as the fit statistic,[20] the maximum-likelihood solution for an ideal model has $\chi^2 = 25,959$. A 3σ deviation from a univariate normal distribution (i.e., for a Gaussian parameter posterior) has a log-likelihood 4.5 lower than the peak. The equivalent range of likelihoods for the given χ^2 distribution is $\chi^2 = [25281.4, 26648.6]$, i.e., $\Delta\chi^2 = [-677.6, -689.6]$ or $\chi^2_{\rm red} = 0.974, 1.026$. However, the best-fit model was only able to achieve $\chi^2 = 1.677e3$ for the r-band image, or $\chi^2_{\rm red} = 6.46$. An increase in χ^2 of just 10.6 produces a $\Delta LL = -4.5$, so the range of acceptable χ^2 shrinks by a factor of approximately $\chi^2_{\rm red}$. Accordingly, the posterior parameter PDFs shrink by a comparable margin, depending on the linearity of the model. For a large number of degrees of freedom, this behavior approaches that of a normal distribution.

Finally, we note that overfitting is strongly disfavored by the χ^2 statistic—a desirable feature for data with reliable errors, as in images dominated by shot noise from a large number of

---

[20] We ignore model parameters in the effective degrees of freedom, as the number of model parameters in nonlinear models is poorly defined (Andrae et al. 2010).

counts. For lower signal-to-noise ratio and/or where other terms like read noise are important, Poisson statistics or the so-called Cash (1979) statistic should be used.

Appendix E
Thick Disk Surface Brightness Profile Fits

As discussed in Section 4, the best-fit MagRite model has an unusually large disk scale height of 1.67 kpc. We ran a number of tests by modifying ProFit to fit a thick disk with a $sech^2$ vertical profile using a similar integration scheme as described in Section 3.2. We superpose 30 Sérsic disks above and below the disk midplane by shifting the profile center along the minor axis, weighting each disk by the total mass within each vertical bin. This is not the most efficient integration method for a 3D density profile and has limited accuracy for highly inclined thick disks, but it is analogous to the MagRite method and ideal for model comparisons.

The results of fitting thick disk profiles to G79635's r-band image are shown in Figure 11. First, fitting a thick exponential disk (second row/column) significantly improves the residuals over a thin exponential disk. The improvement is seen largely in the two underdense regions between the galaxy center and spiral arms (roughly NNE and SSW from the galaxy center). However, the thin Sérsic disk achieves similar improvements without significantly worsening the residuals anywhere else in the disk. Thus, it is clear that a thick disk can compensate for deviations from a pure exponential disk, but not necessarily as well as by simply modifying the radial profile. Unfortunately, the best-fit ProFit disk thicknesses are unrealistically large—4.83 kpc for the exponential disk and 2.35 for the Sérsic disk.

Figure 12 shows the posterior distribution from the ProFit r-band fits to mock data using the same input parameters. The "Thick" model shows the same chains as in Figure 6, where the mock data were generated with a more plausible scale height $z_d = 0.1R_{e,d}/1.67835$ (equivalent to $0.1R_d$ for an exponential disk), and with the other best-fit parameters taken from the best thin disk fit. A second fit was run on mock data with clipped residuals added back in. Specifically, after generating the PSF-convolved model image, we add $\chi\sigma \tanh(abs(\chi/1.25))^{30}$, where σ is the per-pixel uncertainty. The tanh scaling smoothly truncates residuals below 2σ, so the disk more closely follows a Sérsic profile. Because overdensities like spiral arms and star-forming regions tend to be more significant than underdensities, the residuals have a small net positive flux of slightly under one percent of the disk luminosity, which is reflected in a small positive bias in disk luminosity in Figure 12 compared to the input parameters. The size and Sérsic index are somewhat biased, but the scale height and axis ratio are significantly over- and underestimated, respectively, indicating that structured residuals with a small net flux can severely bias poorly constrained and/or degenerate parameters even when the model is a reasonable approximation to the data.

The fact that the scale height and inclination are highly degenerate is not surprising—if the disk's vertical density profile is the same as its radial profile, then the scale height and inclination will be completely degenerate. Using a $sech^2$ vertical density profile rather than exponential limits does not prevent this degeneracy. MagRite achieves a tighter constraint on the inclination by fitting the velocity map. Of course, the mass model also modifies the rotation curve, but
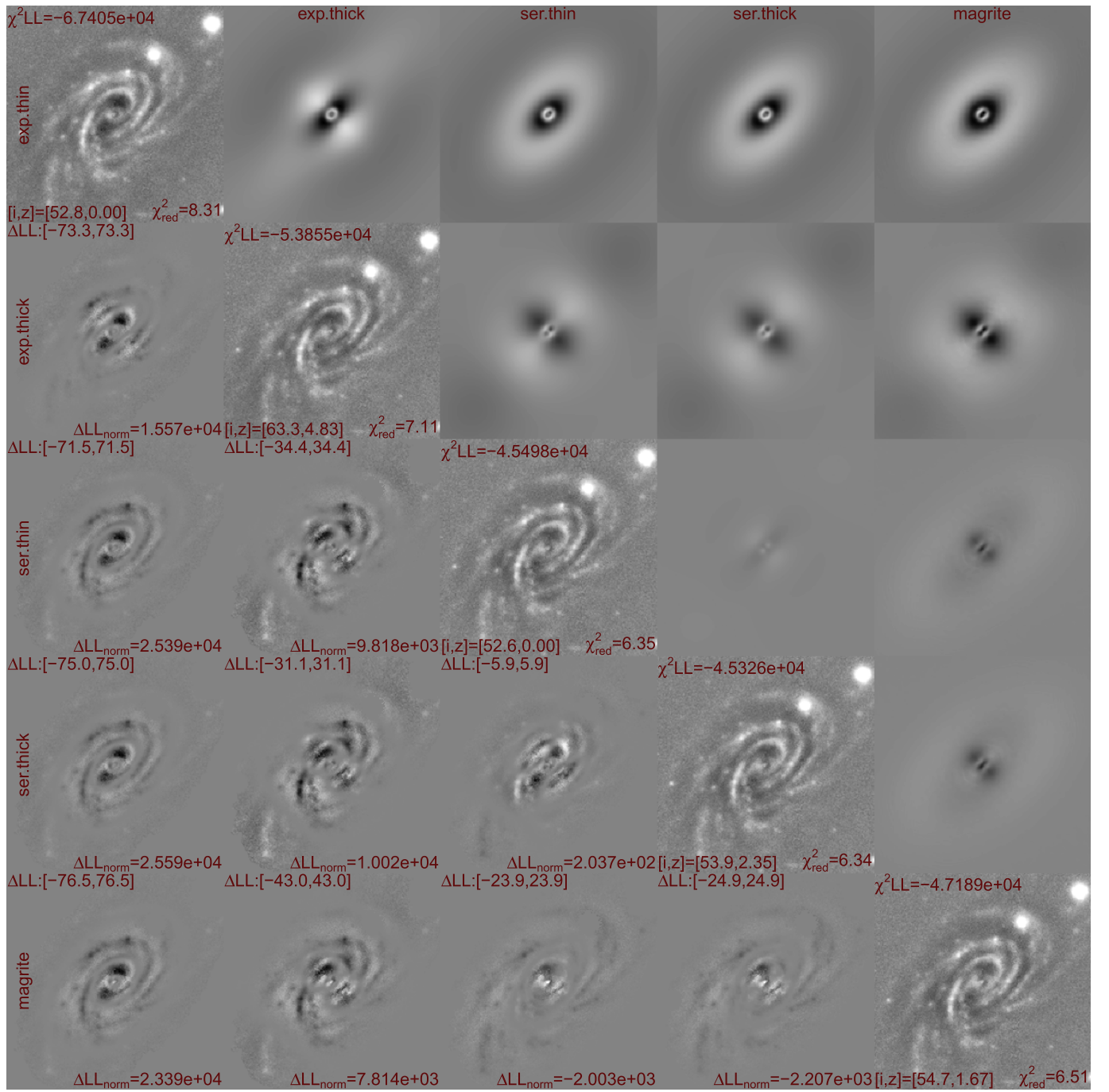
Figure 11. Comparison of r-band residuals for five G79635 model fits. The models include four ProFit models—thin exponential disk ("exp.thin"), thick exponential disk ("exp.thick"), thin Sérsic disk ("ser.thin"), and thick Sérsic disk ("ser.thick")—as well as the MagRite fit. Diagonal panels show the model residuals $\chi = (\text{data-model})/\sigma$ on a common scale (from approximately $-22 < \chi < 22$), where dark colors correspond to excess in the model and light colors to excess in the data. Also listed are the log-likelihoods, assuming a χ^2 distribution; $\chi^2_{\rm red}$; and the disk inclination i and scale height in kpc z. Panels above the diagonal show the differences between models on a common linear scale centered on zero; dark colors are where the model in a given row is brighter than the model in the given column. Panels below the diagonal show the differences in the model residuals (χ) on arbitrary linear scales, again centered on zero and where dark colors represent a better fit in the model in the given row than in the model in the given column. The range of changes in the per-pixel log-likelihood (assuming normally distributed errors) is show in the top left, and the total change in the normal log-likelihood is listed at the bottom right. As expected, residuals improve with added model complexity, but are slightly worse in MagRite (which also fits the kinematics and g-band image) than in the ProFit Sérsic thick disk fit; however, both ProFit thick disk models yield excessively large scale heights.

the stellar mass is independently constrained by the flux maps. Unfortunately, the kinematic constraints on the disk scale height itself are weak. In principle, the disk dispersion is related to the disk's vertical structure, but this is (mostly) independently parametrized in GalactICS by σ_{R0}, and SAMI's spectral resolution is not fine enough to measure typical disk dispersions anyway. Thus, there are insufficient data to guarantee an accurate best-fit scale height, and a strong prior based on observations of edge-on disks should be used in practice.
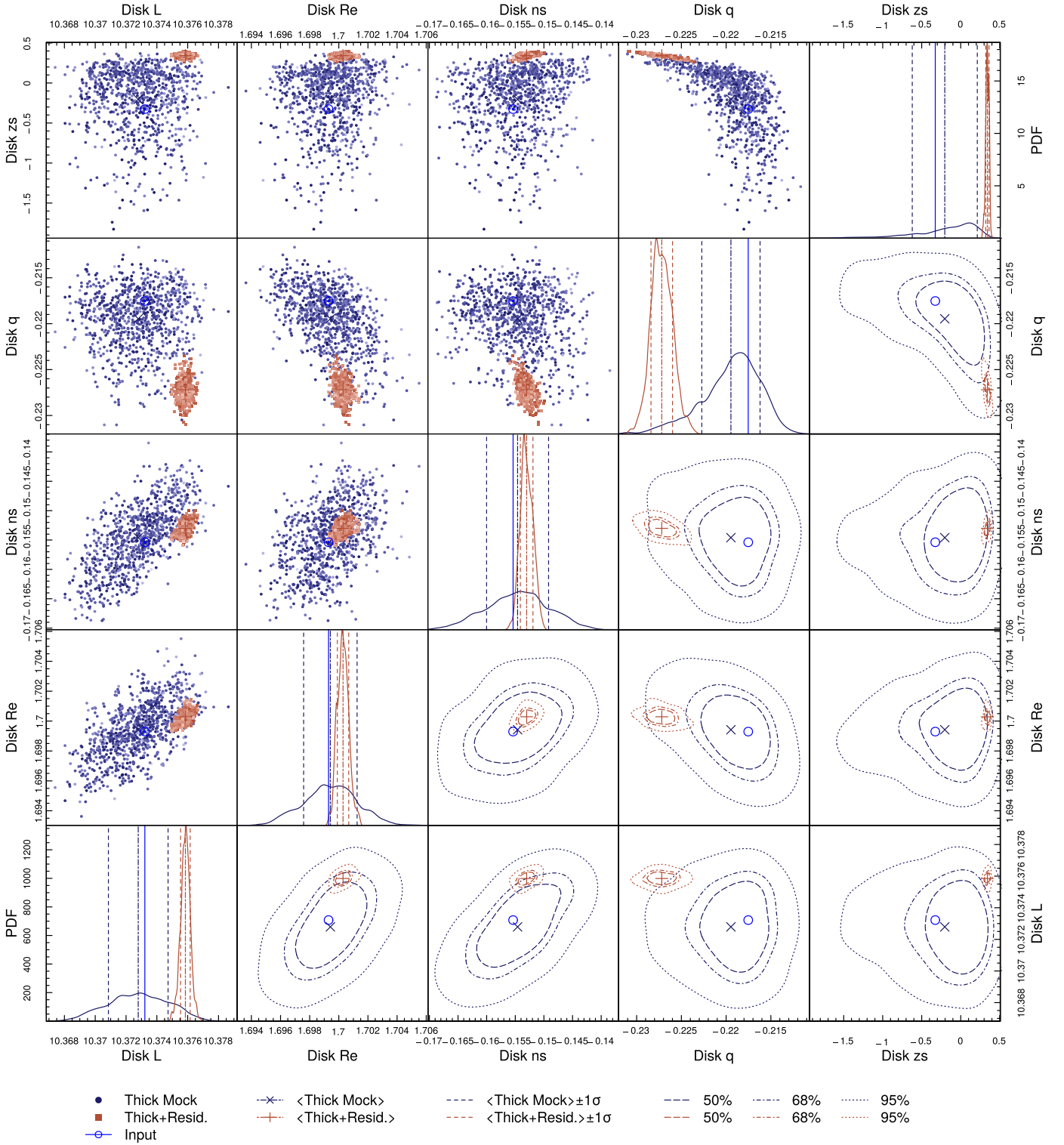
Figure 12. Triangle plot showing joint posterior parameter distributions ($L \equiv \log 10(L_r/L_\odot)$, $Re \equiv \log 10(R_e/\mathrm{kpc})$, $n \equiv \log 10(n_s)$, $zs \equiv \log 10(z_d/\mathrm{kpc})$, and where q is the disk axis ratio) for the ProFit$r$-band Sérsic disk fits to mock data, with and without including clipped residuals from the best-fit model. Panels are structured as in Figure 6. The fit to the mock data with clipped residuals ("Thick+Resid.") has smaller uncertainties and is also significantly biased, particularly for the scale height and inclination. The magnitude bias is due to the small net positive flux of the residuals. For clarity, points in the upper-left quadrant are thinned by factors of 5 and 10 for the "Thick Mock" and "Thick+Resid." samples, respectively.

Appendix F
Summary of the GalactICS Method

We provide a brief description of the methods used by GalactICS to generate DFs for composite galaxy models containing any number of disk-like and spherical components composed of stars, dark matter, and gas (Kuijken & Dubinski 1995; Widrow & Dubinski 2005; Widrow

et al. 2008). We use the Binney & Tremaine (2008) convention below for denoting the cylindrical radius as uppercase R and spherical radius as lowercase r, as well as the physics/ISO 80000 convention of θ for the polar angle and ψ for the azimuthal angle (in contrast with Section 3.2. We also use the relative binding energy $\mathcal{E} \equiv E$ and relative potential $\Psi \equiv -\Phi$, where $\mathcal{E} = \Psi - v^2/2$ and v is the velocity (such that $v^2/2$ is the specific kinetic energy).

A galaxy model is defined parametrically by the mass profiles for each collisionless component. Stellar bulges and dark-matter halos are defined by spherical models described by a radial profile $\rho(r)$. For example, we use the 3D deprojected Sérsic profile (Prugniel & Simien 1997) to describe the bulge, which is given by

$$\rho_b(r) = \rho_0 \left(\frac{r}{R_e} \right)^{-p} \exp[-b_n(r/R_e)^{1/n_s}], \qquad (1)$$

where the parameters are the characteristic density ρ_0, the projected half-mass radius R_e, and the the Sérsic index n_s. The two parameters b_n and p are structural quantities depending on n_s. They are well-approximated by the formulae

$$p = 1 - 0.6097/n_s + 0.05563/n_s^2 \qquad (2)$$
$$b = 2n_s - 1/3 + 0.009876/n_s \qquad (3)$$

for $0.6 < n_s < 10$ and $10^{-2} < R/R_e < 10^3$ (Terzić & Graham 2005).

There are numerous options for halo profiles depending on one's theoretical bias, including profiles with constant density cores or power-law cusps (Merritt et al. 2006). In this paper, we use a double power-law model to describe the halo,

$$\rho_h(r) = \rho_s \left(\frac{r}{r_s} \right)^{-\alpha} \left(1 + \frac{r}{r_s} \right)^{\alpha - \beta}, \qquad (4)$$

where $(\alpha, \beta) = (1, 3)$ is the NFW profile and $(\alpha, \beta) = (1, 4)$ is the Hernquist (1990) profile; such double power-law models are sometimes referred to as generalized NFW or Hernquist profiles.

Disks are flat axisymmetric models and are approximated by the density law

$$\rho_d(R, z) = \frac{M_d}{4\pi R_d^2 z_d} \exp(-R/R_d) \operatorname{sech}^2(z/z_d) \qquad (5)$$

(van der Kruit & Searle 1981). We emphasize that this is not the exact disk density but rather a close approximation to the final density law derived in the computation of the disk distribution function defined below.

Finally, we force the density of each component to smoothly approach zero by multiplying each profile by a truncation function. We truncate density laws using a logistic function defined by

$$T(t) = (1 + e^t)^{-1}, \qquad (6)$$

with

$$t = \frac{r - r_t}{\delta r_t}, \qquad (7)$$

where r_t is the truncation radius and δr_t is the radial width of the truncation interval. Equation (6) is a simple representation of a smooth step function chosen for its computational efficiency and continuous derivatives.

The method computes an axisymmetric DF for the system of the form

$$f(\mathcal{E}, L_z, \mathcal{E}_z) = f_d(\mathcal{E}, L_z, \mathcal{E}_z) + f_b(\mathcal{E}) + f_h(\mathcal{E}), \qquad (8)$$

where $\mathcal{E}$ is the relative binding energy $\mathcal{E} \equiv -E$, L_z is the z-component of the angular momentum, and $\mathcal{E}_z$ is the z energy defined below. The bulge and halo are functions of energy alone and so are modeled as spherical isotropic systems. The disk DF is defined as a function of three integrals of motion. The first two are the usual energy and z-component of angular momentum for axisymmetric systems but we introduce a third approximate integral $\mathcal{E}_z = \Psi_z - v_z^2/2$, where Ψ_z is the vertical potential defined as $\Psi_z \equiv \Psi(R, z) - \Psi(R, z = 0)$, where Ψ is the relative gravitational potential of the system $\Psi = -\Phi$.

With these various definitions in hand, we can describe a numerical procedure for computing the component DFs. First, consider a purely spherical system composed of multiple components—we will consider modifications when including a thin disk component later. The construction of an isotropic DF $f(\mathcal{E})$ from a potential-density pair can be accomplished using Eddington's formula (e.g., Binney & Tremaine 2008),

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi^2} \left[\int_0^{\mathcal{E}} \frac{d\Psi}{\sqrt{\mathcal{E} - \Psi}} \frac{d^2\rho}{d\Psi^2} + \frac{1}{\sqrt{\mathcal{E}}} \left(\frac{d\rho}{d\Psi} \right)_{\Psi=0} \right]. \qquad (9)$$

For a system of total density $\rho(r)$, we can compute the total potential using the integral expression

$$\Psi(r) = 4\pi G \left[\frac{1}{r} \int_0^r dr' r'^2 \rho(r') + \int_r^\infty dr' r' \rho(r') \right]. \qquad (10)$$

To use the Eddington formula, one needs to determine the function $\rho(\Psi)$ and its derivatives up to second order. In general, it is difficult to find an analytic solution so we use the following numerical method. We first define a grid with n radial positions equally spaced in logarithmic space defined by

$$u_i = \log r_i/r_0, \qquad (11)$$

where r_0 is a reference radius and r_i is the grid point radius for $i = 1..n$. With this transformation, we can compute the potential at the grid positions u_i as

$$\Psi(u_i) = 4\pi G r_0^2 \left[e^{-u_i} \int_{-\infty}^{u_i} du' e^{3u'} \rho(u') \right.$$
$$\left. + \int_{u_i}^\infty du' e^{2u'} \rho(u') \right]. \qquad (12)$$

In practice, the infinite limits for the inner and outer integrals can be replaced with the initial and final points of the logarithmic grid u_1 and u_n without loss of accuracy. The inner integral is just the mass versus radius, and this becomes insignificant if the innermost radius is sufficiently small. For the outer integral, since the density drops to zero at a finite radius defined by the truncation function, there is no need to integrate beyond this point. Accurate and stable numerical solutions are achievable with 200 grid spacings per dex, making the new method much faster.

The integral in Eddington's formula can be solved numerically by creating a tabulated function of the density ρ for each of the model functions versus the total Ψ on the logarithmic radial grid. We use the interpolation modules in the GNU science library (GSL) to create a splined function for

$\rho(\Psi)$ and its first and second derivatives. We can then numerically solve the integral in Eddington's formula to determine a DF for each spherical component independently. The DF is also determined as a tabulated function by finding f for each value of $\mathcal{E} = \Psi$ on the radial grid.

To incorporate the disk component in this scheme, we use the ad hoc method introduced by Widrow & Dubinski (2005). The disk density law of equation is spherically averaged to create an additional spherical density component that is part of the total potential. The individual DFs computed from the scheme above then include a good approximation of the disk potential in their derivation. We see below that when we include the flattened disk, the spherical density profiles of the bulge and halo that are consistent with their DF's are modified slightly from the ideal spherical case but remain close to the original definition.

We now consider the construction of the disk DF. Kuijken & Dubinski (1995, hereafter KD95) introduced this DF to describe the disk:

$$f_d(E_p, L_z, E_z) = \frac{\Omega(R_c)}{(2\pi)^{3/2}\kappa(R_c)}\frac{\widetilde{\rho}_d(R_c)}{\widetilde{\sigma}_R^2(R_c)\widetilde{\sigma}_z(R_c)}$$
$$\times \exp\left[-\frac{E_p - E_c(R_c)}{\widetilde{\sigma}_R^2(R_c)} - \frac{E_z}{\widetilde{\sigma}_z^2(R_c)}\right], \quad (13)$$

where $E_p = E - E_z$ is the energy in planar motions, L_z is the z-component of the angular momentum, R_c and E_c are the radius and energy of the circular orbit with angular momentum L_z, and Ω and κ are the circular orbital and radial frequencies derived from the total potential. As shown in KD95, one can obtain the density by integrating over velocities, and the resulting density in the plane is $\widetilde{\rho}_d(R)$ with fractional errors $O(\widetilde{\sigma}_R^2/v_c^2)$.

The disk density can be obtained by integrating the disk DF over the velocities, and the result generates a midplane disk density equal to $\widetilde{\rho}_d(R)$ plus fractional terms of $O(\widetilde{\sigma}_R^2/v_c^2)$. By construction, this disk DF works best for cool, thin disks where the epicyclic approximation is valid for disk star orbits, although warmer and thicker disks are still good equilibria in practice (see Appendix B).

The goal is to find a set of "tilde" functions in the disk DF, $\widetilde{\rho}$, $\widetilde{\sigma}_R$, and $\widetilde{\sigma}_z$, that closely approximates the disk density in Equation (5). To this end, we use the density law

$$\rho_d(R, z) = \frac{M_d}{4\pi R_d^2 z_d}e^{-R/R_d}\exp\left[C_1\frac{\Psi_z(R, z)}{\Psi_z(R, C_0 z_d)}\right]$$
$$\times T\left(\frac{r - r_t}{\delta r_t}\right). \quad (14)$$

The constants C_0 and $C_1 = \ln\mathrm{sech}^2(C_0)$ are chosen so that the run of the vertical density of the disk at a given radius R approximates $\mathrm{sech}^2(z/z_d)$. In practice, a good choice is $C_0 = 3$, corresponding to the equivalence of the vertical density at three scale heights for the target density of Equation (5) and this disk density.

At this point, we have built a DF for each component with the density defined in terms of the total potential $\Psi(R, z)$. To achieve self-consistency, one needs to solve for the total potential,

$$\nabla^2\Psi(R, z) = -4\pi G[\rho_d(R, \Psi, \Psi_z) + \rho_b(\Psi) + \rho_h(\Psi)]. \quad (15)$$

As in KD95, we use the iterative method of Prendergast & Tomer (1970) to find a numerical solution. The method proceeds by making an initial guess of the potential Ψ, computing the densities on the right side of Equation (15), and then re-solving for Ψ. This process is iterated until Ψ relaxes to a solution. In practice, we use a multipole expansion of Ψ in spherical coordinates $(r, \cos\theta)$ with the radius defined on the logarithmic grid. For an axisymmetric system, we can write the potential on the logarithmic radial grid as the multipole expansion

$$\Psi(r, \theta) = 4\pi G r_0^2 \sum_{l=0}^{\infty}\frac{P_l(\cos\theta)}{2l + 1}$$
$$\times \left(e^{-(l+1)u}\int_{-\infty}^{u}du\ e^{(l+3)u}a_l(u)\right.$$
$$\left.+ e^{lu}\int_{u}^{\infty}du\ e^{(2-l)u}a_l(u)\right), \quad (16)$$

where $u = \log(r/r_0)$ and the functions $a_l(u)$ are given by

$$a_l(u) = \int \sin\theta\ d\theta P_l(\cos\theta)\rho(u, \theta) \quad (17)$$

(e.g., Binney & Tremaine 2008). We can replace the infinite limits by the end points of the grid as before without losing accuracy. We use convergence of the "tidal" radius of the total density within some error tolerance to stop the iteration. The tidal radius is just the finite radius where the total density drops to zero for the model. The series is truncated for some l sufficiently large to approximate the flattened potential.

We describe some modifications to this procedure that speed up convergence and overall accuracy. The initial guess to the total potential is the composite spherical potential derived in the first stage. We first perform an iterative sequence to monopole order until convergence. In a second phase, we restart the iteration with this solution, gradually adding the terms for higher-order expansions. Furthermore, the new potential derived at each iteration is determined as a weighted average of the newly determined potential and the previous one. In general, convergence to a high order in l can be reached in a few tens of iterations.

Following KD95, we also use an analytic "high harmonics" disk potential to improve the accuracy of the multipole expansion at a lower order in l. This analytic potential is

$$\Psi_d^{\dagger} = \frac{GM_d z_d}{2R_d^2}\ln\cosh(z/z_d)e^{-r/R_d}T\left(\frac{r - r_t}{\delta r_t}\right), \quad (18)$$

where one notes that the radial parameter is the spherical radius. One can use the identity

$$\nabla^2 f(r)\ln\cosh(z) = f''(r)\ln\cosh(z)$$
$$+ \frac{2f'(r)}{r}[z\tanh(z) + \ln\cosh(z)]$$
$$+ f(r)\mathrm{sech}^2(z) \quad (19)$$

to derive an analytic disk density $\rho_d^{\dagger}$ from $\Psi_d^{\dagger}$. The last term reproduces a $\mathrm{sech}^2(z)$ disk to $O(z/R)^2$ while the other terms are generally small for thin disks and tend to zero at the origin. In the iterative method described above, we replace the disk density with the residual $\delta\rho_d = \rho_d - \rho_d^{\dagger}$ when determining the total potential from the multipole expansion. The total final

19

potential is computed as the sum of the multipole expansion plus $\Psi_d^\dagger$. In practice, this method allows an accurate representation of the thin disk model for expansion orders of $l < 10$, greatly speeding up the overall procedure.

The final step involves the finding "tilde" functions in the disk DF that match the disk density in Equation (14). The first function $\widetilde{\sigma}_R^2(R)$ can be set arbitrarily, but following KD95, we use the observationally inspired profile $\widetilde{\sigma}_R^2 = \sigma_{R,0}^2 \exp(-R/R_d)$, where the central radial velocity dispersion $\sigma_{R,0}^2$ is a free parameter for the disk. Normally, this parameter is chosen so that the disk is stable, i.e., Toomre $Q > 1$ across the radial extent of the disk. The remaining functions—the midplane density $\widetilde{\rho}_d(R)$ and the vertical velocity dispersion $\widetilde{\sigma}_z^2$—are iteratively adjusted for each radius on the grid such that the midplane density and the density at $z = z_d$ are the same as that in Equation (14). Finally, the tilde functions can be represented numerically with splines, and thus the disk DF is fully specified.

In summary, the final products of this moderately complex procedure are fully specified potentials and densities for each component defined by multipole expansions with modifications for the thin disk to improve accuracy for lower-order l. Each component also has a well-defined DF determined by Eddington's method for spherical components and the constructed disk DF from Equation (13). The multipole coefficients of the potential expansion are tabulated on the logarithmic grid and represented as splined functions in the code, allowing rapid evaluation of the potential and density at any point. The spherical DFs as $f(\mathcal{E})$ and the tilde functions are computed at the predefined grid coordinates. The code takes tens of seconds on current hardware to tabulate these functions, depending on the maximum multipole order, making it practical to use for fitting galaxy observations.

ORCID iDs

D. S. Taranu ⓘ https://orcid.org/0000-0001-6268-1882
L. M. R. Fogarty ⓘ https://orcid.org/0000-0002-9601-468X
J. van de Sande ⓘ https://orcid.org/0000-0003-2552-0021
B. Catinella ⓘ https://orcid.org/0000-0002-7625-562X
L. Cortese ⓘ https://orcid.org/0000-0002-7422-9823
A. S. G. Robotham ⓘ https://orcid.org/0000-0003-0429-3579
J. T. Allen ⓘ https://orcid.org/0000-0002-2092-8768
J. Bland-Hawthorn ⓘ https://orcid.org/0000-0001-7516-4016
M. Colless ⓘ https://orcid.org/0000-0001-9552-8075
R. L. Davies ⓘ https://orcid.org/0000-0001-7897-3812
M. J. Drinkwater ⓘ https://orcid.org/0000-0003-4867-0022
S. P. Driver ⓘ https://orcid.org/0000-0001-9491-7327
Á. R. López-Sánchez ⓘ https://orcid.org/0000-0001-8083-8046
A. M. Medling ⓘ https://orcid.org/0000-0001-7421-2944
J. R. Mould ⓘ https://orcid.org/0000-0003-3820-1740
C. Power ⓘ https://orcid.org/0000-0002-4003-0904
S. N. Richards ⓘ https://orcid.org/0000-0002-5368-0068

References

Aihara, H., Armstrong, R., Bickerton, S., et al. 2017, arXiv:1702.08449
Andrae, R., Schulze-Hartung, T., & Melchior, P. 2010, arXiv:1012.3754
Bernstein, G. M., & Jarvis, M. 2002, AJ, 123, 583
Binney, J., & McMillan, P. 2011, MNRAS, 413, 1889
Binney, J., & Tremaine, S. 2008, Galactic Dynamics (2nd ed.; Princeton, NJ: Princeton University Press)
Brooks, A. M., Papastergis, E., Christensen, C. R., et al. 2017, arXiv:1701.07835
Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2857
Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Cash, W. 1979, ApJ, 228, 939
Catinella, B., Haynes, M. P., & Giovanelli, R. 2007, AJ, 134, 334
Cecil, G., Fogarty, L. M. R., Richards, S., et al. 2016, MNRAS, 456, 1299
Ciardullo, R., Durrell, P. R., Laychak, M. B., et al. 2004, ApJ, 614, 167
Courteau, S., Cappellari, M., de Jong, R. S., et al. 2014, RvMP, 86, 47
Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872
de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, A&A, 582, A62
de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, ExA, 35, 25
de Vaucouleurs, G. 1959, HDP, 53, 275
de Zeeuw, P. T., Bureau, M., Emsellem, E., et al. 2002, MNRAS, 329, 513
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971
Dubinski, J. 1996, NewA, 1, 133
Fogarty, L. M. R., Scott, N., Owers, M. S., et al. 2015, MNRAS, 454, 2050
Freeman, K. C. 1970, ApJ, 160, 811
Gadotti, D. A. 2009, MNRAS, 393, 1531
Galassi, M. 2009, A GNU Manual, GNU Scientific Library: Reference Manual (3rd ed.; Bristol: Network Theory), xvi
Graham, A. W., & Driver, S. P. 2005, PASA, 22, 118
Hansen, N. 2006, in Towards a New Evolutionary Computation. Advances on Estimation of Distribution Algorithms, ed. J. Lozano et al. (Berlin: Springer) 75–102
Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, AJ, 142, 170
Hernquist, L. 1990, ApJ, 356, 359
Johnston, E. J., Häußler, B., Aragón-Salamanca, A., et al. 2017, MNRAS, 465, 2317
Keller, S. C., Schmidt, B. P., Bessell, M. S., et al. 2007, PASA, 24, 1
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, MNRAS, 439, 1245
Kuijken, K., & Dubinski, J. 1995, MNRAS, 277, 1341
Law, D. R., Yan, R., Bershady, M. A., et al. 2015, AJ, 150, 19
Maraston, C., & Strömbäck, G. 2011, MNRAS, 418, 2785
Martinsson, T. P. K., Verheijen, M. A. W., Westfall, K. B., et al. 2013, A&A, 557, A130
Merritt, D., Graham, A. W., Moore, B., Diemand, J., & Terzić, B. 2006, AJ, 132, 2685
Moffat, A. F. J. 1969, A&A, 3, 455
Muñoz-Mateos, J. C., Sheth, K., Gil de Paz, A., et al. 2013, ApJ, 771, 59
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Obreschkow, D., & Glazebrook, K. 2014, ApJ, 784, 26
Pastrav, B. A., Popescu, C. C., Tuffs, R. J., & Sansom, A. E. 2013, A&A, 557, A137
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Portail, M., Gerhard, O., Wegg, C., & Ness, M. 2017, MNRAS, 465, 1621
Prendergast, K. H., & Tomer, E. 1970, AJ, 75, 674
Puglielli, D., Widrow, L. M., & Courteau, S. 2010, ApJ, 715, 1152
R Core Team 2016, R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing)
Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951
Robotham, A. S. G., Taranu, D. S., Tobar, R., Moffett, A., & Driver, S. P. 2017, MNRAS, 466, 1513
Romanowsky, A. J., & Fall, S. M. 2012, ApJS, 203, 17
Sánchez, S. F., García-Benito, R., Zibetti, S., et al. 2016, A&A, 594, A36
Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, A&A, 538, A8
Sánchez-Janssen, R., Ferrarese, L., MacArthur, L. A., et al. 2016, ApJ, 820, 69
Schwarzschild, M. 1979, ApJ, 232, 236
Sérsic, J. L. 1968, Atlas de Galaxias Australes (Cordoba: Observatorio Astronomico, Universidad Nacional de Cordoba)
Sharp, R., Allen, J. T., Fogarty, L. M. R., et al. 2015, MNRAS, 446, 1551
Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11
Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223

Tabor, M., Merrifield, M., Aragón-Salamanca, A., et al. 2017, MNRAS, 466, 2024
Taranu, D. S., Dubinski, J., & Yee, H. K. C. 2013, ApJ, 778, 61
Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587
Terzić, B., & Graham, A. W. 2005, MNRAS, 362, 197
Toomre, A. 1964, ApJ, 139, 1217
Trick, W. H., Bovy, J., & Rix, H.-W. 2016, ApJ, 830, 97
van de Sande, J., Bland-Hawthorn, J., Fogarty, L. M. R., et al. 2017, ApJ, 835, 104

van der Kruit, P. C., & Searle, L. 1981, A&A, 95, 105
Vasiliev, E., & Athanassoula, E. 2015, MNRAS, 450, 2842
Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, AJ, 136, 2563
Wang, J., Koribalski, B. S., Serra, P., et al. 2016, MNRAS, 460, 2143
Widrow, L. M., & Dubinski, J. 2005, ApJ, 631, 838
Widrow, L. M., Pym, B., & Dubinski, J. 2008, ApJ, 679, 1239
Wright, A. H., Robotham, A. S. G., Driver, S. P., et al. 2017, MNRAS, 470, 283